DTU Library

Sampling Problems with Spectral Coherence

Kristensen, Leif; Kirkegaard, Peter

Publication date:
1986

Document Version
Publisher's PDF, also known as Version of record

Risø-R-526

RISØ

Risø-R-526

Sampling Problems with Spectral Coherence

Leif Kristensen and Peter Kirkegaard

Risø National Laboratory, DK-4000 Roskilde, Denmark
February 1986

RISØ-R-526

SAMPLING PROBLEMS WITH SPECTRAL COHERENCE

Leif Kristensen and Peter Kirkegaard

Abstract. This report deals with the confidence of the spectral coherence derived from a measured time series by a discrete Fourier transform. The estimated coherence will be based on a block averaging procedure, characterized by a certain number of degrees of freedom. A discussion is given of the significance of this number, and how many degrees of freedom are necessary to get a trustworthy sampling of the coherence. The study is carried out by utilizing the statistical results available for complex Gaussian product sums (Goodman statistics).

6th February, 1986

Risø National Laboratory, DK 4000 Roskilde, Denmark

ISBN 87-550-1193-4
ISSN 0106-2840

Grafisk Service, Risø, 1986

CONTENTS

1. INTRODUCTION

A common tool for the study of random stationary time series is the so-called spectral coherence or just coherence as discussed in many textbooks, e.g. Lumley and Panofsky (1964) and Panofsky and Dutton (1984). Following their definition the spectral coherence

$$\mathrm{coh}(\omega) = \frac{|\phi_{xy}(\omega)|^2}{\phi_{xx}(\omega)\ \phi_{yy}(\omega)} \tag{1}$$

is given by the absolute square of the cross-spectrum $\phi_{xy}(\omega)$ between the two time series x(t) and y(t), divided by the product of their power spectra. By this definition coh(ω) is always greater than or equal to zero and by Schwarz's inequality less than or equal to one. Alternatively, the coherence has been defined as the square root of the right side of (1) (Goodman, 1957). We are adhering to the first definition in the following.

In practical applications when dealing with geophysical time series, only one or a few realizations can usually be obtained and since $\phi_{xy}(\omega)$ is defined as ensemble averages over infinitely many realizations there is a statistical uncertainty in the experimentally determined spectral coherence. This is brought out rather dramatically if one tries to compute the coherence by using a discrete Fourier transform on only one realization. In this case the coherence becomes identically one for all frequences ω (as rediscovered the hard way by many inexperienced scientists, including the authors). With one realization available one can subdivide the record in a number of shorter records and in this way obtain an ensemble with a finite number of realizations M. In general the estimated coherence, which in any case is bounded by zero and one, will attain values different from one. Alternatively, one can block-average the spectral estimates $\phi_{xy}(\omega)$ from one realization. We will show that by averaging M estimates around each frequency we get what is roughly equivalent to a subdivision of the original record in M subrecords. Both techniques are applied and the product of the number of subdivisions and the number of averaged spectral estimates is often considered the so-called number of degrees of freedom. If special windows or weighting functions are used the definition is not so simple. Here we stay away from these complications and refer to the literature (Bendat and Piersol 1971, Koopmans 1974, Amos and Koopmans 1963). However, we shall include a discussion about how the finite record length modifies our simple concept and how it is possible to define an effective number of degrees of freedom.

From this discussion we conclude that M = 1 is certainly insufficient for making a coherence estimate. How large must M then be? Is two good enough? or ten? or should M = 100? It is the purpose here to bring to attention the work by Goodman (1957) on distributions of spectral parameters in Gaussian processes, which enables us to answer the question: How large must M be in order to obtain a given statistical confidence in $coh(\omega)$? We do not want to let M be larger than absolutely necessary since an increase in M for a given record length means a decrease in spectral resolution and an increase in spectral distortion.

2. BASIC CONSIDERATIONS

In this and the following section we want to go in some detail in showing that the theory of Goodman's can be applied to spectral coherences. We shall demonstrate that the procedures we use to calculate spectra by themselves secure that the criteria for the validity of the theory are satisfied.

Let $x(t)$ and $y(t)$ be two ergodic time series with ensemble means equal to zero. Since an ergodic time series is stationary there is no loss of generality in assuming that the ensemble means are zero. We imagine that $x(t)$ and $y(t)$ are defined for all times, but only sampled over the time T. In order to set the stage we first assume that x(t) and y(t) are known for all t in the interval $[-T/2, T/2]$. We define the Fourier amplitudes by

$$\hat{x}(\omega;T) = \frac{1}{2\pi} \int_{-T/2}^{T/2} x(t)\, e^{i\omega t}\, dt \tag{2a}$$

and

$$\hat{y}(\omega;T) = \frac{1}{2\pi} \int_{-T/2}^{T/2} y(t)\, e^{i\omega t}\, dt \quad , \tag{2b}$$

and the covariance between these amplitudes by

$$E_{xy}(\omega',\omega'';T) = \langle\, \hat{x}(\omega';T)\ \hat{y}^*(\omega'';T)\, \rangle \quad , \tag{3}$$

where the brackets mean ensemble averaging, superstars complex conjugation, and where we have used that $\langle x(t)\rangle = \langle y(t)\rangle = 0$ implies $\langle\hat{x}(\omega;T)\rangle = \langle\hat{y}(\omega;T)\rangle = 0$.

By substituting (2a) and (2b) in (3) we obtain

$$E_{xy}(\omega',\omega'';T) = \frac{1}{(2\pi)^2} \int_{-T/2}^{T/2} dt' \int_{-T/2}^{T/2} dt''\, R_{xy}(t',t'')\, e^{-i(\omega''t''-\omega't')} \tag{4}$$

where

$$R_{xy}(t',t'') = \langle x(t')y(t'')\rangle \tag{5}$$

is the cross-covariance function. Since the time series x(t) and y(t) are stationary $R_{xy}(t',t'')$ depends on t' and t" only through the difference t" - t' and we may write

$$R_{xy}(t',t'') = R_{xy}(0,t'' - t') \equiv R_{xy}(t''-t') \quad , \tag{6}$$

suppressing the first argument for convenience. By introducing the variable transform

$$\left.\begin{aligned} t &= (t' + t'')/2 \\ \tau &= t'' - t' \end{aligned}\right\} \tag{7}$$

(4) becomes

$$E_{xy}(\omega',\omega'';T) = \frac{1}{(2\pi)^2} \int_{-T}^{T} R_{xy}(\tau) e^{-i\frac{\omega'+\omega''}{2}\tau} d\tau \int_{-\frac{T-|\tau|}{2}}^{\frac{T-|\tau|}{2}} e^{-i(\omega''-\omega')t}\, dt \tag{8}$$

or

$$E_{xy}(\omega',\omega'';T) = \frac{1}{(2\pi)^2} \int_{-T}^{T} \frac{\sin\left(\frac{\omega''-\omega'}{2}[T-|\tau|]\right)}{\frac{\omega''-\omega'}{2}} R_{xy}(\tau)\, e^{-i\frac{\omega'+\omega''}{2}\tau} d\tau\ . \tag{9}$$

In the limit $T\to\infty$ this equation becomes

$$\lim_{T\to\infty} E_{xy}(\omega',\omega'';T) = \delta(\omega''-\omega')\ \phi_{xy}\left(\frac{\omega'+\omega''}{2}\right) , \tag{10}$$

where

$$\delta(x) = \frac{1}{\pi} \lim_{K\to\infty} \frac{\sin Kx}{x} \tag{11}$$

is the Dirac delta function and

$$\phi_{xy}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} R_{xy}(\tau)\ e^{-i\omega\tau}\, d\tau \tag{12}$$

the cross-spectrum of x(t) and y(t). By the Fourier inversion rule we obtain from (12)

$$R_{xy}(\tau) = \int_{-\infty}^{\infty} \phi_{xy}(\omega)\ e^{i\omega\tau}\, d\omega\ . \tag{13}$$

Equation (10) states that Fourier amplitudes of different frequencies are uncorrelated in the limit $T \to \infty$; thus it seems natural to assume that for "large enough" values of T we need not be concerned with $E_{xy}(\omega', \omega''; T)$ for $\omega' \neq \omega''$. For $\omega' = \omega'' = \omega$ (9) reads

$$E_{xy}(\omega, \omega; T) = \frac{1}{(2\pi)^2} \int_{-T}^{T} (T - |\tau|)\, R_{xy}(\tau) e^{-i\omega\tau}\, d\tau$$

$$= \frac{1}{\Delta\omega} \frac{1}{2\pi} \int_{-T}^{T} \left(1 - \frac{|\tau|}{T}\right) R_{xy}(\tau)\, e^{-i\omega\tau}\, d\tau \quad , \tag{14}$$

where

$$\Delta\omega = \frac{2\pi}{T} \tag{15}$$

is the smallest frequency that can be resolved by a Fourier analysis of the time series x(t) and y(t) over the finite time of duration T. If T is large enough we see, by comparing (12) and (14), that

$$\Psi_{xy}(\omega; T) \equiv E_{xy}(\omega, \omega; T)\, \Delta\omega \approx \phi_{xy}(\omega) \quad . \tag{16}$$

We note that

$$\int_{-\infty}^{\infty} \Psi_{xy}(\omega; T)\, d\omega = R_{xy}(0) = \int_{-\infty}^{\infty} \phi_{xy}(\omega)\, d\omega \tag{17}$$

by use of (11) and (13) and conclude that $\Psi_{xy}(\omega;T)$ and $\phi_{xy}(\omega)$ cover the same area, equal to the total covariance $R_{xy}(0)$. In a loose way we can say that $\Psi_{xy}(\omega;T)$ is a distorted estimate of $\phi_{xy}(\omega)$, in which covariance is moved from low to high frequencies without loss of total covariance.

Another observation is that (3) and (16) state that for x(t) = y(t), $\Psi_{xy}(\omega;T) = \Psi_{xx}(\omega;T)$ is never negative. Since this is true for all values of T we conclude that the power spectrum $\phi_{xx}(\omega)$ is real and non-negative.

With these basic concepts it is now possible to introduce in a natural way the effects of the discrete, rather than continuous, sampling and Fourier analysis. The invention of the Fast Fourier Transform (FFT) by Cooley and Tukey (1965) made it more efficient to Fourier-transform the time series themselves rather than the covariance functions in the calculation of the spectra. This is of course history now, but let us recapitulate the line of arguments that justifies the common approach to spectral analysis by use of digital computers.

The two time series x(t) and y(t) are sampled over the period T at times separated by the increment Δt. The total number of sampling times is N and we have

$$T = N\Delta t \tag{18}$$

Instead of (2a) and (2b) we now write

$$\hat{x}[k;N] = \frac{1}{N} \sum_{\ell=0}^{N-1} x[\ell]\, e^{2\pi i \ell k/N} \tag{19a}$$

and

$$\hat{y}[k;N] = \frac{1}{N} \sum_{\ell=0}^{N-1} y[\ell]\, e^{2\pi i \ell k/N} \quad , \tag{19b}$$

where

$$\left.\begin{aligned} x[\ell] &= x(\ell\Delta t) \\ y[\ell] &= y(\ell\Delta t) \end{aligned}\right\} \ \ell = 0,1,2,\ldots, N-1 \quad . \tag{20}$$

On comparing (20) with (16) we infer that

$$\Psi_{xy}[k;N] \equiv \langle \hat{x}[k;N]\; \hat{y}^{*}[k;N] \rangle \tag{21}$$

will be closely related to the spectrum $\phi_{xy}(\omega)$ at the frequency

$$\omega = \omega_k = k\Delta\omega = k\,\frac{2\pi}{T} \quad . \tag{22}$$

If we substitute (19a) and (19b) in (21) and use (20) and (5) and (6) we get

$$\Psi_{xy}[k;N] = \frac{1}{N^2} \sum_{\ell'=0}^{N-1} e^{2\pi i \ell' k/N} \sum_{\ell''=0}^{N-1} e^{-2\pi i \ell'' k/N} R_{xy}(\Delta t(\ell''-\ell')). \tag{23}$$

We want to obtain a relation between $\Psi_{xy}[k;N]$ and $\phi_{xy}(\omega_k)$ and therefore we use (13) to eliminate $R_{xy}(\Delta t(\ell''-\ell'))$ in (23). Rearranging the order of summations and integration leads to the following expression

$$\Psi_{xy}[k;N] = \int_{-\infty}^{\infty} \frac{\sin^2\{\frac{\omega_k-\omega}{2}T\}}{N^2\sin^2\{\frac{\omega_k-\omega}{2N}T\}} \phi_{xy}(\omega)\, d\omega \quad , \tag{24}$$

where ω_k is given by (22). In general, (24) is a rather complicated convolution integral, but fortunately N is usually so large that we can use the approximation

$$\frac{\sin^2\{\frac{\omega_k-\omega}{2}T\}}{N^2\sin^2\{\frac{\omega_k-\omega}{2N}T\}} \simeq \frac{2\pi}{T} \sum_{m=-\infty}^{\infty} \delta(\frac{2\pi}{\Delta t} m+\omega-\omega_k) \tag{25}$$

Introducing the so-called Nyquist frequency

$$\omega_N = \pi/\Delta t \tag{26}$$

and making use of (15), (24) may be written in the form

$$\frac{\Psi_{xy}[k;N]}{\Delta\omega} = \sum_{m=-\infty}^{\infty} \phi_{xy}(\omega_k + 2m\omega_N) \ . \tag{27}$$

The left-hand side constitutes what we hope to be a good approximation to the spectral density at frequency ω_k, obtained by a discrete finite Fourier (DFT) technique, and the right-hand side is an infinite sum of true spectral densities separated by $2\omega_N$. If the time series contains most of its turbulent energy at low frequencies, i.e. frequencies absolutely smaller than ω_N, then the m = 0 term is dominant. However, higher terms often cannot be neglected. This "spectral contamination" is called spectral aliasing. We see that $\Psi_{xy}[k;N]$ is periodic in k with the period N, corresponding to the increment $2\pi/\Delta t$ in ω_k. Further, we see from (19) and (21) that $\Psi_{xy}[-k;N] = \Psi^*_{xy}[k;N]$. Therefore, we need only be interested in $\Psi_{xy}[k;N]$ in the interval (0,N/2). It is possible by suitable low-pass filtering to suppress the effect of spectral aliasing. In the following section we shall assume that disregarding spectral aliasing is justified. We can then concentrate on the problem that in geophysics we never have an ensemble of infinitely many realizations; as a rule we have only one.

3. SMOOTHING

The ensemble averaging in (21) is a mathematical idealization that does not correspond to reality in geophysics. Instead we form from our single realization

$$\chi_{xy}[k;N;1] = \hat{x}[k;N]\,\hat{y}^*[k;N] \quad . \tag{28}$$

In order to improve the statistical confidence we average $\chi_{xy}[k;N;1]$ over a number M of consecutive values:

$$\chi_{xy}[k;N;M] = \frac{1}{M}\sum_{m=0}^{M-1} \chi_{xy}[k+m;N;1] \quad . \tag{29}$$

This is called smoothing. We shall investigate how good an approximation $\chi_{xy}[k;N;M]$ is to $\Psi_{xy}[k;N]$ given by (21). First we note that

$$\langle\chi_{xy}[k;N;M]\rangle = \frac{1}{M}\sum_{m=0}^{M-1} \Psi_{xy}[k+m;N] \approx \Psi_{xy}\left[k+\frac{M-1}{2};N\right] \tag{30}$$

(in this context we allow the first argument of χ_{xy} to be non-integral with an obvious meaning).

Thus $\chi_{xy}[k;N;M]$ can naturally be interpreted as an estimator of Ψ_{xy} evaluated at a point halfway between k and k+(M-1). If Ψ_{xy}'s variation with k is linear over this interval, the estimator is unbiased. Otherwise the averaging introduces a systematical error or bias B roughly proportional to M^2 and to the second derivative of ϕ_{xy}:

$$B \equiv \frac{1}{M}\sum_{m=0}^{M-1} \Psi_{xy}[k+m;N] - \Psi_{xy}[k+\frac{M-1}{2};N] \simeq \frac{\Delta\omega}{24}(M\Delta\omega)^2\ \phi''_{xy}(\omega_k). \quad (31)$$

In addition, there is an error of a statistical nature associated with the fluctuation of $\chi_{xy}[k;N;M]$ around its ensemble mean $\langle\chi_{xy}[k;N;M]\rangle$. The variance of this error is

$$\sigma^2_{xy}[k;N;M] = \langle|\chi_{xy}[k;N;M] - \langle\chi_{xy}[k;N;M]\rangle|^2\rangle \quad . \quad (32)$$

A small value of this quantity means a precise estimator, and if the bias is negligible we may expect $\chi_{xy}[k;N;M]$ to be close to $\Psi_{xy}[k+(M-1)/2;N]$. In the following we shall make an approximate calculation of the error variance.

Substitution of (29) and (28) in (32) yields

$$\sigma^2_{xy}[k;N;M] =$$

$$\frac{1}{M^2}\sum_{m'=0}^{M-1}\sum_{m''=0}^{M-1} \{\langle\hat{x}[k+m';N]\hat{y}^*[k+m';N]\hat{x}^*[k+m'';N]\hat{y}[k+m'';N]\rangle$$
$$-\langle\hat{x}[k+m';N]\hat{y}^*[k+m';N]\rangle\langle\hat{x}^*[k+m'';N]\hat{y}[k+m'';N]\rangle\} \quad . \quad (33)$$

We see that fourth-order moments of Fourier amplitudes are involved and this usually will make the task intractable. However, if the Fourier amplitudes are joint Gaussian distributed then fourth-order moments can be expressed by second-order moments. This is probably the case to a good approximation, because each amplitude is a complex, weighted sum of a large number N of random numbers as (19a) and (19b) show; according to the central limit theorem the amplitudes will be asymtotically Gaussian for $N\to\infty$. Then the well-known Isserlis relation for four joint-Gaussian real variables A, B, C and D with zero means,

$$\langle ABCD\rangle = \langle AB\rangle\langle CD\rangle + \langle AC\rangle\langle BD\rangle + \langle AD\rangle\langle BC\rangle, \tag{34}$$

generalized to the case of arbitrary complex joint Gaussians A, B, C and D (Koopmans, 1974) can be applied. Using (34), (33) reduces to

$$\sigma^2{}_{xy}[k,N,M] = \frac{1}{M^2}\sum_{m'=0}^{M-1}\sum_{m''=0}^{M-1}\{E_{xx}[k+m';k+m'';N]\ E^*_{yy}[k+m';k+m'';N]$$

$$+ E_{xy}[k+m',-k-m'';N]\ E^*_{yx}[k+m',-k-m'';N]\} \tag{35}$$

where

$$E_{xy}[k',k'';N] = \langle\hat{x}[k';N]\hat{y}^*[k'';N]\rangle \quad . \tag{36}$$

We get, by substitution of (19a) and (19b),

$$E_{xy}[k',k'';N] =$$

$$\frac{1}{N^2}\sum_{\ell'=0}^{N-1} e^{2\pi i\ell' k'/N}\sum_{\ell''=0}^{N-1} e^{-2\pi i\ell'' k''/N}\, R_{xy}(\Delta t(\ell''-\ell')), \tag{37}$$

and by inserting (13) we obtain, almost as in case of (24),

$$E_{xy}[k',k'';N] =$$

$$(-1)^{k''-k'}\, e^{i\pi\frac{k''-k'}{N}}\int_{-\infty}^{\infty}\phi_{xy}(\omega)\,\frac{\sin(\pi k'-\omega T/2)}{N\sin\left(\frac{\pi k'-\omega T/2}{N}\right)}\,\frac{\sin(\pi k''-\omega T/2)}{N\sin\left(\frac{\pi k''-\omega T/2}{N}\right)}d\omega. \tag{38}$$

Since we are disregarding spectral aliasing, (38) is well approximated by

$$E_{xy}[k',k'';N] \approx$$

$$(-1)^{k''-k'}\, e^{i\pi\frac{k''-k'}{N}}\int_{-\infty}^{\infty}\phi_{xy}(\omega)\,\frac{\sin(\pi k'-\omega T/2)}{\pi k'-\omega T/2}\,\frac{\sin(\pi k''-\omega T/2)}{\pi k''-\omega T/2}d\omega. \tag{39}$$

We see that (39) is a convolution of the spectrum $\phi_{xy}(\omega)$ with the product of two, rather peaked functions with maxima at $k'\Delta\omega$ and $k''\Delta\omega$. The integrand is therefore significantly different from zero only when $k'\Delta\omega$ and $k''\Delta\omega$ are close to each other. We approximate (39) by

$$E_{xy}[k',k'';N] \approx (-1)^{k''-k'} e^{i\pi\frac{k''-k'}{N}} \phi_{xy}(\Delta\omega\frac{k'+k''}{2}) \times$$

$$\int_{-\infty}^{\infty} \frac{\sin(\pi k'-\omega T/2)}{\pi k'-\omega T/2} \frac{\sin(\pi k''-\omega T/2)}{\pi k''-\omega T/2} d\omega \quad . \tag{40}$$

As shown in Appendix A we find

$$\int_{-\infty}^{\infty} \frac{\sin(\pi k'-\omega T/2)}{(\pi k'-\omega T/2)} \frac{\sin(\pi k''-\omega T/2)}{(\pi k''-\omega T/2)} d\omega = \frac{2}{T} \frac{\sin(\pi(k''-k'))}{k''-k'} \tag{41}$$

so that (40) reduces to

$$E_{xy}[k',k'';N] \approx$$

$$(-1)^{k''-k'} e^{i\pi\frac{k''-k'}{N}} \phi_{xy}(\Delta\omega\frac{k'+k''}{2})\Delta\omega \frac{\sin(\pi(k''-k'))}{\pi(k''-k')} \tag{42}$$

or, since k' and k" are integers,

$$E_{xy}[k',k'';N] = \Delta\omega \; \phi_{xy} \; (\Delta\omega\frac{k'+k''}{2}) \; \delta_{k'k''} \; . \tag{43}$$

The expression (35) reduces to

$$\sigma^2_{xy}[k;N;M] \approx \frac{1}{M^2} \sum_{m=0}^{M-1} (\Delta\omega)^2 \; \phi_{xx}(\Delta\omega(k+m)) \; \phi^*_{yy}(\Delta\omega(k+m)) \tag{44}$$

since the second term gives zero contribution for $k > 0$. We can reduce the right-hand side even further by assuming that the spectra ϕ_{xx} and ϕ_{yy} vary insignificantly over the range $M\Delta\omega$:

$$\sigma_{xy}^2[k;N;M] \simeq (\Delta\omega)^2 \frac{\phi_{xx}(\omega_k + \frac{M-1}{2}\Delta\omega)\ \phi^*_{yy}(\omega_k + \frac{M-1}{2}\Delta\omega)}{M}$$

$$\simeq (\Delta\omega)^2 \frac{\phi_{xx}(\omega_k)\ \phi^*_{yy}(\omega_k)}{M} \quad . \tag{45}$$

We see that in this approximation the error variance is inversely proportional to the number of estimates entering the averaging (29), and we may write

$$\sigma_{xy}^2[k;N;M] \simeq \sigma^2_{xy}[k;N;1]/M \ . \tag{46}$$

The value of M becomes equal to the number of degrees of freedom and $\sigma_{xy}[k;N;M]$ and $\sigma_{xy}[k;N;1]$ are the standard errors for M and one degree of freedom, respectively. This interpretation is equivalent to the statement that neighbour spectral estimates are statistically independent. However, we expected that, depending on the length of the record T, there would be some statistical dependence between neighbour spectral estimates. This dependence would be assumed to disappear as T increases without limit. It turns out that in going from (39) to (40), where we take the spectrum outside the integral, we are making an approximation that is slightly incorrect. In other words, the statistical dependence is hidden in the difference between the right-hand sides of (39) and (40). We shall return to this rather difficult subject in the next section. The opposite dependencies of the bias (31) and the error variance (45) on M render the practical choice of this parameter a matter of compromise.

We conclude this section by showing that block averaging of M spectral estimates is roughly equivalent to subdivision of the original record of length N numbers in M subrecords of length L = N/M numbers. This subdivision is illustrated in Fig. 1.

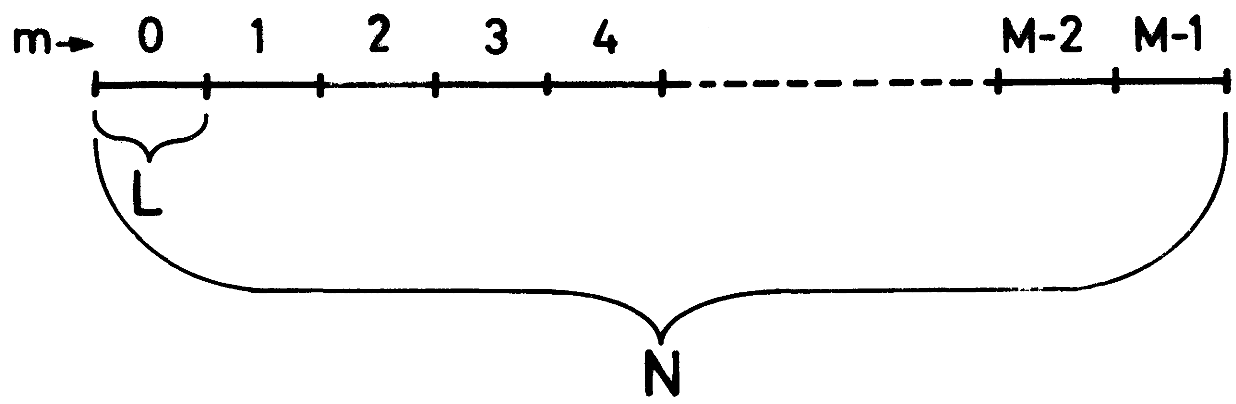


Fig. 1. Illustration of the record subdivision in M records, each of length L.

For convenience we repeat (19a) and (19b)

$$\hat{x}[k;N] = \frac{1}{N} \sum_{n=0}^{N-1} x[n]\, e^{2\pi ik\, n/N} \tag{19a}$$

$$k = (0,1,\ldots,\ N-1) \bmod N$$

$$\hat{y}[k;N] = \frac{1}{N} \sum_{n=0}^{N-1} y[n]\, e^{2\pi ik\, n/N} \tag{19b}$$

Their inverse transformations are

$$x[n] = \sum_{k=0}^{N-1} \hat{x}[k;N]\, e^{-2\pi ik\, n/N} \tag{47a}$$

$$y[n] = \sum_{k=0}^{N-1} \hat{y}[k;N]\, e^{-2\pi ik\, n/N} \tag{47b}$$

From $\hat{x}[k;N]$ and $\hat{y}[k;N]$ we first form an estimate of the spectral density at frequency $\omega_k = k\, 2\pi/N\Delta t$

$$\chi_{xy}(\omega_k;T) = \chi_{xy}[k;N;1]/\Delta\omega = \frac{N\Delta t}{2\pi}\, \hat{x}[k;N]\hat{y}^*[k;N]. \tag{48}$$

For each of the M subintervals we write, in analogy to (19a) and (19b)

$$\left.\begin{aligned}\hat{x}[k;m;L] &= \frac{1}{L}\sum_{\ell=0}^{L-1} x[mL+\ell]\, e^{2\pi ik\ell/L} \qquad &(49a)\\ \hat{y}[k;m;L] &= \frac{1}{L}\sum_{\ell=0}^{L-1} y[mL+\ell]\, e^{2\pi ik\ell/L} \qquad &(49b)\end{aligned}\right\}\quad \begin{aligned}&m = 0,1\ldots,\ M-1;\\ &k = (0,1,\ldots,L-1) \bmod L\end{aligned}$$

For a particular value m we define

$$\chi_{xy}(\omega_k;T/M;m) = \frac{L\Delta t}{2\pi}\,\hat{x}[k;m;L]\hat{y}^*[k;m;L] \tag{50}$$

and the average spectral estimate becomes

$$\chi_{xy}(\omega_k;T/M) = \frac{1}{M}\sum_{m=0}^{M-1} \chi_{xy}(\omega_k;T/M;m), \tag{51}$$

where now $\omega_k = kM\Delta\omega$. We want to show that this averaging is roughly equivalent to block averaging of $\chi_{xy}(\omega_k;T)$ over M neighbouring values of k. We substitute (50) and (49) in (51) and use (47) and (48) to obtain

$$\chi_{xy}(\omega_k;T/M) = \frac{1}{M}\sum_{m=0}^{M-1}\frac{L\Delta t}{2\pi}\,\hat{x}[k;m;L]\hat{y}^*[k;m;L]$$

$$= \sum_{k'=0}^{N-1}\sum_{k''=0}^{N-1} \hat{x}[k';N]\hat{y}^*[k'';N] \times$$

$$\frac{\Delta t}{2\pi L}\sum_{\ell'=0}^{L-1} e^{2\pi i\ell'(\frac{k}{L}-\frac{k'}{N})}\sum_{\ell''=0}^{L-1} e^{-2\pi i\ell''(\frac{k}{L}-\frac{k''}{N})}\,\frac{1}{M}\sum_{m=0}^{M-1} e^{2\pi i(k''-k')\frac{m}{M}}$$

$$= \sum_{k'=0}^{N-1} \hat{x}[k';N]\hat{y}^*[k';N]\;\frac{\Delta t}{2\pi L}\Big|\sum_{\ell=0}^{L-1} e^{2\pi i\left(k-\frac{k'}{M}\right)\frac{\ell}{L}}\Big|^2$$

$$= \sum_{k'=0}^{N-1} \chi_{xy}(\omega_{k'};T)\;\frac{1}{M}\;\frac{\sin^2(\pi(k-k'/M))}{L^2\sin^2(\pi(k-k'/M)/L)}\quad . \tag{52}$$

We see that $\chi_{xy}(\omega_k;T/M)$ is a non-uniformly weighted mean of $\chi_{xy}(\omega_k;T)$, where the "width" is about M "raw spectral estimates". This is so becauce $\chi_{xy}(\omega_k;T)$ is periodic in k with the period N, so that (52) also can be rewritten in the form

$$\chi_{xy}(\omega_k;T/M) = \sum_{k'=-N/2}^{N/2-1} \chi_{xy}\left(\omega_{k-\frac{k'}{M}};T\right) \frac{1}{M} \frac{\sin^2\left(\pi \frac{k'}{M}\right)}{L^2\sin^2\left(\pi \frac{k'}{M}\frac{1}{L}\right)} \tag{53}$$

with an even weighting function, and because the norm of this weighting function is unity, i.e.

$$\sum_{k=-N/2}^{N/2-1} \frac{1}{M} \frac{\sin^2\left(\pi \frac{k}{M}\right)}{L^2\sin^2\left(\pi \frac{k}{M}\frac{1}{L}\right)} = 1, \tag{54}$$

an equation which follows easily from the last rewriting in (52).

4. EFFECTIVE DEGREES OF FREEDOM

In the preceding section we pointed out that we expected the following: If we consider the number of degrees of freedom assigned to a smoothed spectral estimate equal to the arithmetic mean of M raw spectral estimates from one realization of finite length T, it then would be less than M because the raw spectral estimates are statistically independent only in the limit $T\to\infty$. This was based on our considerations about the error variance (32). We will now discuss this problem in more detail; however, this will take us on a detour from the direct road to the main goal, namely to obtain knowledge about the statistics of sampled coherence. Since we are going to neglect the difference between ference between M and the degrees of freedom anyway in the following sections, we suggest that the reader skip the present section in the first perusal.

It has not been possible for us to derive a general relation between the number of degrees of freedom M_{eff}, corrected for the finite duration of the time series, and the number of raw estimates M in the smoothing process. However, we have gained some insight by studying a particular case where $x(t) = y(t)$, and the spectrum has the specific form (Cauchy)

$$\phi_{xx}(\omega) = \phi(\omega) = \frac{\sigma^2\mathcal{T}}{\pi}\,\frac{1}{1+(\omega\mathcal{T})^2}, \tag{55}$$

corresponding to a first-order Markov process. Here σ^2 is the ensemble variance of the time series $x(t)$ and

$$\mathcal{T} = \int_0^\infty \frac{R_{xx}(\tau)}{\sigma^2}\,d\tau = \pi\,\frac{\phi(0)}{\sigma^2} \tag{56}$$

the integral scale. Substituting (55) in (39) we obtain

$$E_{xx}[k',k'';N] = E[k',k'';N] = (-1)^{k''-k'}\,e^{i\pi\frac{k''-k'}{N}}\,\frac{\sigma^2\mathcal{T}}{\pi}\times$$

$$\int_{-\infty}^{\infty}\frac{1}{1+(\omega\mathcal{T})^2}\;\frac{\sin\left(\pi k'-\frac{\omega T}{2}\right)}{\pi k'-\frac{\omega T}{2}}\;\frac{\sin\left(\pi k''-\frac{\omega T}{2}\right)}{\pi k''-\frac{\omega T}{2}}\,d\omega \tag{57}$$

Introducing

$$s = \omega T/2 \tag{58}$$

as integration parameter and

$$\theta = T/(2\mathcal{T}) \tag{59}$$

as a dimensionless measure of the duration of the time series, we can write

$$E[k',k";N] = (-1)^{k"-k'} e^{i\pi\frac{k"-k'}{N}} \frac{\sigma^2}{\pi}\frac{1}{\theta} \times$$

$$\int_{-\infty}^{\infty} \frac{\sin(\pi k'-s)}{\pi k'-s} \frac{\sin(\pi k"-s)}{\pi k"-s} \frac{ds}{1+(s/\theta)^2} \tag{60}$$

It is shown in Appendix A that the integral in (60) can be evaluated analytically, and we can rewrite (60) in the following way:

$$E[k',k";N] =$$

$$\frac{\sigma^2(-1)^{k"-k'} e^{i\pi\frac{k"-k'}{N}}}{2(\theta^2+\pi^2k'^2)(\theta^2+\pi^2k"^2)} \Big\{ \theta(2\theta^2+\pi^2k'^2+\pi^2k"^2) \frac{\sin(\pi(k"-k'))}{\pi(k"-k')}$$
$$- (\theta^2-\pi^2k'k") \cos(\pi(k"-k'))$$
$$+ e^{-2\theta}((\theta^2-\pi^2k'k")\cos(\pi(k'+k"))-\theta\pi(k'+k")\sin(\pi(k'+k"))) \Big\}. \tag{61}$$

Since k' and k" are integers we get

$$E[k',k";N] =$$

$$\sigma^2 \Big\{ \frac{\theta}{\theta^2+\pi^2k'^2} \delta_{k'k"} - \frac{1}{2} e^{i\pi\frac{k"-k'}{N}} \frac{\theta^2-\pi^2k'k"}{(\theta^2+\pi^2k'^2)(\theta^2+\pi^2k"^2)} (1-e^{-2\theta}) \Big\}. \tag{62}$$

According to (35) we must now evaluate the sum of two terms

$$A = E_{xx}[k',k";N]\ E^*_{yy}[k',k";N] = |E[k',k";N]|^2 \tag{63}$$

and

$$B = E_{xy}[k',-k";N]\ E^*_{yx}[k',-k";N] = |E[k',-k";N]|^2 , \tag{64}$$

where

$$\begin{Bmatrix} k' \\ k'' \end{Bmatrix} = \begin{Bmatrix} k + m' \\ k + m'' \end{Bmatrix} . \tag{65}$$

Without loss of generality we can assume that $k > 0$ in these considerations and we get

$$A = \sigma^4 \left\{ \frac{\theta^2}{(\theta^2+\pi^2k'^2)^2} \delta_{k'k''} - (1-e^{-2\theta}) \frac{\theta(\theta^2-\pi^2k'^2)}{(\theta^2+\pi^2k'^2)^3} \delta_{k'k''} \right.$$

$$\left. + \frac{1}{4} (1-e^{-2\theta})^2 \frac{(\theta^2-\pi^2k'k'')^2}{(\theta^2+\pi^2k'^2)^2 (\theta^2+\pi^2k''^2)^2} \right\} \tag{66}$$

and

$$B = \sigma^4 \left\{ \frac{\theta^2}{(\theta^2+\pi^2k'^2)^2} \delta_{k',-k''} - (1-e^{-2\theta}) \frac{\theta(\theta^2-\pi^2k'^2)}{(\theta^2+\pi^2k'^2)^3} \delta_{k',-k''} \right.$$

$$\left. + \frac{1}{4} (1-e^{-2\theta})^2 \frac{(\theta^2+\pi^2k'k'')^2}{(\theta^2+\pi^2k'^2)^2(\theta^2+\pi^2k''^2)^2} \right\}$$

$$= \frac{1}{4} \sigma^2(1-e^{-2\theta})^2 \frac{(\theta^2+\pi^2k'k'')^2}{(\theta^2+\pi^2k'^2)^2 (\theta^2+\pi^2k''^2)^2} , \tag{67}$$

where in (67) there is only one term since both k' and k'' are assumed positive. We want to find the sum of (66) and (67) and we write it in the form

$$A+B = \sigma^4 \left\{ \frac{\theta(\theta-1+e^{-2\theta})}{(\theta^2+\pi^2k'^2)^2} \delta_{k'k''} + 2\theta(1-e^{-2\theta}) \frac{\pi^2k'^2}{(\theta^2+\pi^2k'^2)^3} \delta_{k'k''} \right.$$

$$\left. + \frac{1}{2} (1-e^{-2\theta})^2 \frac{\theta^4+\pi^2k'^2\pi^2k''^2}{(\theta^2+\pi^2k'^2)^2 (\theta^2+\pi^2k''^2)^2} \right\} . \tag{68}$$

Substituting in (35) we get

$$\sigma_{xy}^2[k;N;M] = \sigma^2[k;N;M] =$$

$$\frac{\sigma^4}{M^2}\,\theta(\theta-1+e^{-2\theta}) \sum_{m=0}^{M-1} \frac{1}{(\theta^2+\pi^2(k+m)^2)^2}$$

$$+ \frac{2\sigma^4}{M^2}\,\theta(1-e^{-2\theta}) \sum_{m=0}^{M-1} \frac{\pi^2(k+m)^2}{(\theta^2+\pi^2(k+m)^2)^3}$$

$$+ \frac{\sigma^4}{M^2}\,\frac{\theta^4}{2}\,(1-e^{-2\theta})^2 \left\{ \sum_{m=0}^{M-1} \frac{1}{(\theta^2+\pi^2(k+m)^2)^2} \right\}^2$$

$$+ \frac{\sigma^4}{M^2}\,\frac{1}{2}\,(1-e^{-2\theta})^2 \left\{ \sum_{m=0}^{M-1} \frac{\pi^2(k+m)^2}{(\theta^2+\pi^2(k+m)^2)^2} \right\}^2 , \tag{69}$$

We are considering only cases for which the record length T is much larger than the integral scale $\mathcal{T}$. This means that

$$\theta \gg 1 \tag{70}$$

and since

$$\sum_{m=0}^{M-1} \frac{\pi^2(k+m)^2}{(\theta^2+\pi^2(k+m)^2)^3} < \sum_{m=0}^{M-1} \frac{\theta^2+\pi^2(k+m)^2}{(\theta^2+\pi^2(k+m)^2)^3} = \sum_{m=0}^{M-1} \frac{1}{(\theta^2+\pi^2(k+m)^2)^2} \tag{71}$$

the second term in (69) can be neglected in comparison with the first. The error variance can therefore be well approximated by

$$\sigma^2[k;N;M] \approx \frac{\sigma^4}{M^2} \left\{ \theta^2 \sum_{m=0}^{M-1} \frac{1}{(\theta^2+\pi^2(k+m)^2)^2} \right.$$

$$\left. + \frac{\theta^4}{2} \left[\sum_{m=0}^{M-1} \frac{1}{(\theta^2+\pi^2(k+m)^2)^2} \right]^2 + \frac{1}{2} \left[\sum_{m=0}^{M-1} \frac{\pi^2(k+m)^2}{(\theta^2+\pi^2(k+m)^2)^2} \right]^2 \right\} . \tag{72}$$

Note that πk is not considered small compared to θ in this derivation.

Using (55), (59) and (22) we can now reformulate (72):

$$\sigma^2[k;N;M] \simeq \frac{(\Delta\omega)^2}{M^2} \left\{ \sum_{m=k}^{k+M-1} \phi^2(m\Delta\omega) + \frac{1}{2} \left[\frac{\Delta\omega}{\sigma^2} \sum_{m=k}^{k+M-1} \phi^2(m\Delta\omega) \right]^2 \right.$$

$$\left. + \frac{1}{2} \left[\frac{\Delta\omega}{\sigma^2} \sum_{m=k}^{k+M-1} \phi(m\Delta\omega) \left\{ \phi(0) - \phi(m\Delta\omega) \right\} \right]^2 \right\} . \tag{73}$$

We define

$$\omega_* = \left(k + \frac{M-1}{2}\right) \Delta\omega \tag{74}$$

and obtain the approximate result

$$\sigma^2[k;N;M] \simeq \left\{ \phi(\omega_*)\Delta\omega \right\}^2 \left[\frac{1}{M} + \frac{1}{2} \left\{ \frac{\phi(\omega_*)\Delta\omega}{\sigma^2} \right\}^2 \right.$$

$$\left. + \frac{1}{2} \left\{ \frac{(\phi(0)-\phi(\omega_*))\Delta\omega}{\sigma^2} \right\}^2 \right] . \tag{75}$$

In this equation we recognize the first term as (45). The two following terms are corrections due to the finite record length T. We may recast (75) in the form

$$\sigma^2[k;N;M] = (\phi(\omega_*)\Delta\omega)^2 \left\{ \frac{1}{M} + \lambda(\omega_*) \left(\frac{\mathcal{T}}{T}\right)^2 \right\} \tag{76}$$

with

$$\lambda(\omega_*) = 2 \frac{1+(\omega_*\mathcal{T})^4}{(1+(\omega_*\mathcal{T})^2)^2} . \tag{77}$$

The function $\lambda(\omega_*)$ varies between 1 and 2, such that $\lambda(0) = \lambda(\infty) = 2$, and the minimum 1 is attained for $\omega_* = 1/\mathcal{T}$. Thus for both $\omega_* \to 0$ and $\omega_* \to \infty$ we have

$$\sigma^2[k;N;M] = (\phi(\omega_*)\Delta\omega)^2 \left\{ \frac{1}{M} + 2 \left(\frac{\mathcal{T}}{T}\right)^2 \right\} . \tag{78}$$

If we use the correction term $2(\mathcal{T}/T)^2$ in (78) we make a conservative estimate of the effect of the finite record length. Considering (78) a good estimate of the error variance we have obtained an expression which is just of the form we expected: the error variance is close to being inversely proportional to the number of degrees of freedom M except if the record length is too small. Strictly speaking the result is valid only for power spectra of the Cauchy form (55), and to investigate how generally (78) can be applied to other power spectra, in Appendix B we have carried through a similar analysis for a power spectrum of the form

$$\phi(\omega) = \frac{\sigma^2\mathcal{T}}{\pi} \frac{1}{\{1 + (\frac{\omega\mathcal{T}}{a})^2\}^\alpha} , \qquad (79)$$

where α is a positive constant and $\mathcal{T}$ is given by (56). The constant a is given by the constraint that the integral over $\phi(\omega)$ from $-\infty$ to $+\infty$ is equal to the variance σ^2. A simple calculation shows that

$$a = \sqrt{\pi}\,\frac{\Gamma(\alpha)}{\Gamma(\alpha - \frac{1}{2})} \quad . \qquad (80)$$

It turns out that for $1/2 < \alpha < 1$ (78) is still conservative in the limits of small and large values of ω_* and we therefore suggest this equation be used in general. We can then define an effective number of degrees of freedom M_{eff} by the equation

$$\frac{1}{M_{eff}} = \frac{1}{M} + 2(\frac{\mathcal{T}}{T})^2 \qquad (81)$$

or

$$M_{eff} = \frac{M}{1 + 2(\mathcal{T}/T)^2 M} \qquad (82)$$

and we see the effect of the "principle of diminishing return" when $M \sim (T/\mathcal{T})^2$. In view of the fact that we in general will lose information about the curvature of the spectrum and also introduce a bias by smoothing, we have demonstrated that there is an upper limit to how large we want M to be.

5. GOODMAN DISTRIBUTIONS

Goodman (1957, 1963) has derived a number of important distributions connected with sampling in complex Gaussian processes. Particularly interesting to us among these results is the distribution of the so-called sample coherence of two complex observation sequences, defined by

$$Z^2 = \frac{\left|\frac{1}{n}\sum_{j=1}^{n} X[j]Y^*[j]\right|^2}{\left(\frac{1}{n}\sum_{j=1}^{n}|X[j]|^2\right)\left(\frac{1}{n}\sum_{j=1}^{n}|Y[j]|^2\right)} \tag{83}$$

In this expression $(X[j], Y[j])$, $(j=1,\ldots,n)$ are identically distributed complex Gaussian random variable pairs. In the Cartesian expressions

$$X = X_r + i\ X_i \tag{84a}$$

$$Y = Y_r + i\ Y_i \tag{84b}$$

we shall require that the real random variables X_r, X_i, Y_r, Y_i, are distributed four-variate Gaussian with zero means and a dispersion matrix with the following structure

$$V = \begin{bmatrix} \sigma_x^2 & 0 & \alpha\sigma_x\sigma_y & \beta\sigma_x\sigma_y \\ 0 & \sigma_x^2 & -\beta\sigma_x\sigma_y & \alpha\sigma_x\sigma_y \\ \alpha\sigma_x\sigma_y & -\beta\sigma_x\sigma_y & \sigma_y^2 & 0 \\ \beta\sigma_x\sigma_y & \alpha\sigma_x\sigma_y & 0 & \sigma_y^2 \end{bmatrix} . \tag{85}$$

This equation causes some restrictions on the complex variable pair (X,Y). On the stated assumptions, Goodman (1957) worked out an expression for the probability density function of Z, the square-root coherence:

$$q(z) = \frac{2(1-\gamma^2)^n}{(n-1)!(n-2)!} z(1-z^2)^{n-2} \sum_{k=0}^{\infty} \frac{\gamma^{2k}((n+k-1)!)^2}{(k!)^2} z^{2k} , \tag{86}$$

where

$$\gamma^2 = \frac{|<XY^*>|^2}{<XX^*><YY^*>} \tag{87}$$

is the true coherence. The probability density function p for the quantity $U = Z^2$, which in our terminology is the sample coherence, is easily derived from (86) since

$$p(u) = q(z) \frac{dz}{du} = \frac{1}{2z} q(z) \quad . \tag{88}$$

(Unfortunately, in Goodman's report there is notational confusion between Z and Z^2). The probability density function (86) and many similar distribution results were derived with the aid of the complex Wishart distribution (see Goodman (1963)) which has a relatively simple characteristic function. In (87) "<>" means expectation and is in fact the same operator as the ensemble averaging used in our previous sections. For the variance we shall use the notation Var[ ] here. Note that the expression (87) is the theoretical or "true" coherence between X and Y, in contrast to the estimated or measured coherence given by (83).

A coherence estimator based on the block-averaged spectral estimates given in section 3 is

$$Z^2 = \frac{|\chi_{xy}[k;N;M]|^2}{\chi_{xx}[k;N;M] \ \chi_{yy}[k;N;M]} , \tag{89}$$

valid for the frequency ω_k (or rather ω_*, see (74)). We shall presently show that such an estimator is approximately compatible with (83) - (85). With this proviso we may proceed to draw inferences from the probability density expressions (86) and (88).

In (83) and (86) n plays the role of the number of degrees of freedom, a quantity we denoted M and to which we paid much attention in sections 3 and 4. We shall later compute $\langle Z^2 \rangle$ as well as $\mathrm{Var}[Z^2]$ and find asymptotic formulas for these statistics for large n. Though $\chi_{xy}[k;N;M]$ as stated in (28) - (29) is a block average it may in a first approximation be considered as a simple average over M statistically independent terms $\hat{x}[j;N]\hat{y}^*[j;N]$. In the Cartesian representation of the Fourier amplitudes $\hat{x}[j;N]$ and $\hat{y}[j;N]$:

$$\left.\begin{aligned} \hat{x}[j;N] &= \hat{x}_r + i\,\hat{x}_i \\ \hat{y}[j;N] &= \hat{y}_r + i\,\hat{y}_i \end{aligned}\right\} , \qquad \begin{aligned} &(90a) \\ &(90b) \end{aligned}$$

we must show that the distribution of $(\hat{x}_r,\hat{x}_i,\hat{y}_r,\hat{y}_i)$ is joint Gaussian with a dispersion matrix V of the same structure as (85). The Gaussian property holds approximately for the same reason as given in connection with (33). If in (36) and the approximate relation (43) we let $k' = -k'' = j$ it follows that

$$\langle \hat{x}[j;N]\hat{y}[j;N] \rangle \simeq 0 \tag{91}$$

or, using (90)

$$\langle \hat{x}_r\hat{y}_r - \hat{x}_i\hat{y}_i \rangle + i\langle \hat{x}_r\hat{y}_i + \hat{x}_i\hat{y}_r \rangle \simeq 0 \quad . \tag{92}$$

Consequently,

$$\langle \hat{x}_r\hat{y}_r \rangle = \langle \hat{x}_i\hat{y}_i \rangle \equiv \mu_o \tag{93}$$

and

$$\langle \hat{x}_r\hat{y}_i \rangle = -\langle \hat{x}_i\hat{y}_r \rangle \equiv \mu_1 \tag{94}$$

If we let x = y in (93) we deduce that

$$\langle \hat{x}_r^2 \rangle = \langle \hat{x}_i^2 \rangle \equiv \sigma_x^2 \tag{95}$$

and

$$\langle \hat{y}_r^2 \rangle = \langle \hat{y}_i^2 \rangle \equiv \sigma_y^2 \quad . \tag{96}$$

Similarly, x = y in (94) yields

$$\langle \hat{x}_r \hat{x}_i \rangle = \langle \hat{y}_r \hat{y}_i \rangle = 0 \quad . \tag{97}$$

Now we are equipped to construct the dispersion matrix of $(\hat{x}_r, \hat{x}_i, \hat{y}_r, \hat{y}_i)$. It becomes

$$V = \begin{bmatrix} \sigma_x^2 & 0 & \mu_o & \mu_1 \\ 0 & \sigma_x^2 & -\mu_1 & \mu_o \\ \mu_o & -\mu_1 & \sigma_y^2 & 0 \\ \mu_1 & \mu_o & 0 & \sigma_y^2 \end{bmatrix} . \tag{98}$$

We see that it fulfills the requirements of (85), permitting Goodman's theory to be used.

A natural question to ask at this point is: In what manner does the refined analysis given in section 4, which ended up with the "effective" number of degrees of freedom M_{eff} in (82), affect the validity of the heuristic justification given for the use of Goodman's distribution results? It is not easy to give an exact treatment of this issue. Though M_{eff} was derived from a study of the error variance, it seems very plausible that a good approach will be to take $n = M_{eff}$ in Goodman's formulas, and we believe that this procedure will account for the essential features of coherence sampling. The very fact that (82) need not assign an integral value to M_{eff} precludes an exact treatment; however, we shall assume for simplicity that henceforward $n = M_{eff}$ is a given natural number.

We observe that by (88) and (86) p(u) can be expressed by a hypergeometric function:

$$p(u) = (n-1)(1-\gamma^2)^n(1-u)^{n-2}\ F(n,n;1;\gamma^2 u) \quad . \tag{99}$$

When n = 1 (99) degenerates to $p(u) = \delta(1-u)$, expressing that u = 1 with probability one (cf. remark in the Introduction).

Amos and Koopmans (1963) have devoted much effort in devising sophisticated computational procedures for (86), and they give extensive tables for q(z) for a range of values of z, n, and γ. However, we are more interested in $U = Z^2$ than in Z itself. We notice that by use of the well-known hypergeometric relation

$$F(a,b;c;x) = (1-x)^{c-a-b} F(c-b,c-a;c;x) \tag{100}$$

(99) can be written as a finite sum:

$$p(u) = (n-1) \frac{(1-\gamma^2)^n (1-u)^{n-2}}{(1-\gamma^2 u)^{2n-1}} \sum_{k=0}^{n-1} \frac{(n-1)^2 (n-2)^2 \ldots (n-k)^2}{(k!)^2} \gamma^{2k} u^k . \tag{101}$$

Equation (101) is well suited for a direct calculation of the probability density, at least when n is not excessively large. In this way we computed the graphs in Figure 2, which show p(u) for n = 5, 10, and 50 degrees of freedom, and for the "true" coherence γ^2 = 0.1, 0.5, and 0.9.

We have also investigated the sample phase ϕ_s. This may be defined with reference to (83) as

$$\phi_s = \arg\left(\frac{1}{n} \sum_{j=1}^{n} X[j]Y^*[j]\right) . \tag{102}$$

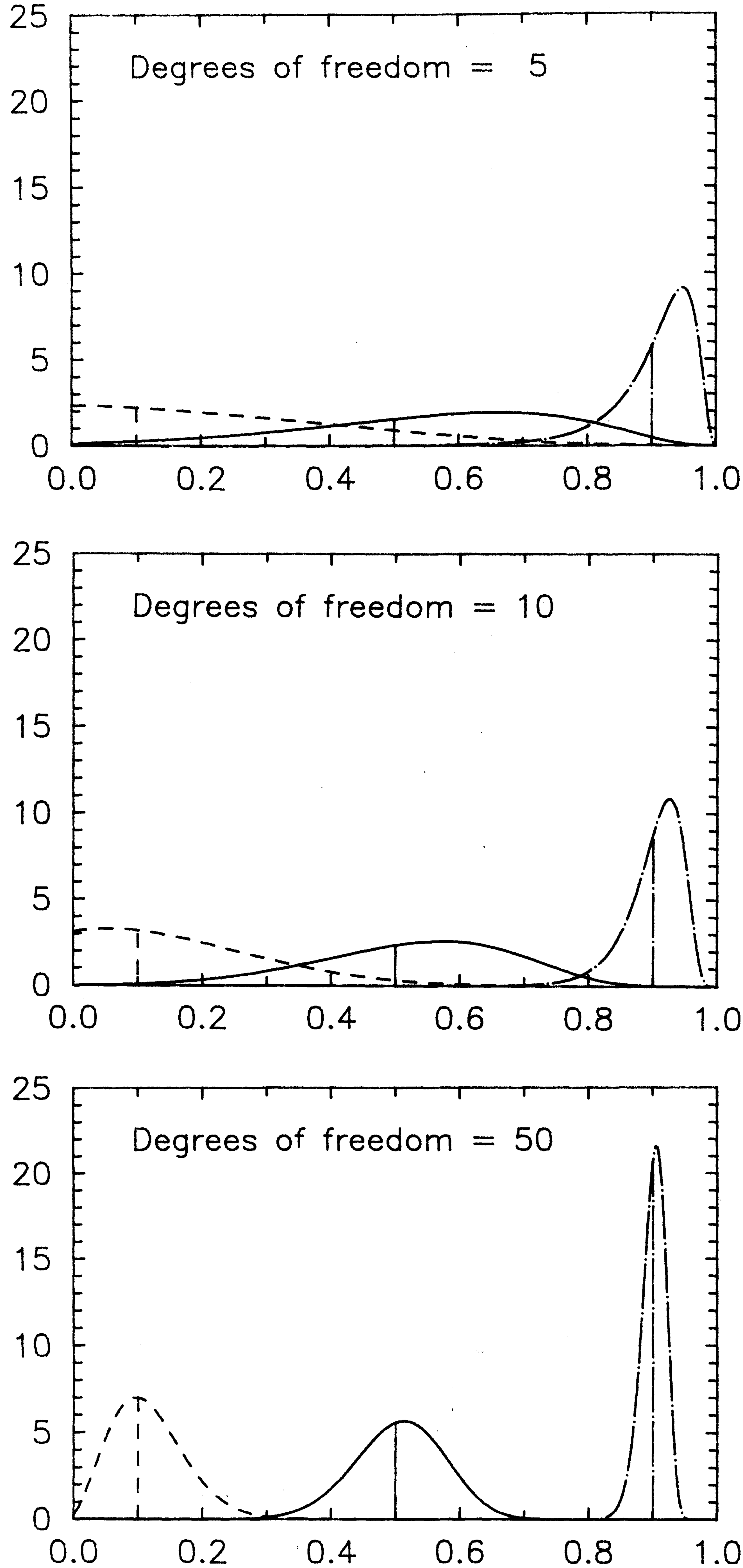
We shall consider the distribution of ϕ_s, or rather its deviation from the "true" phase ϕ_o,

$$\phi \equiv \phi_s - \phi_o , \tag{103}$$

where

$$\phi_o = \arg\langle XY^*\rangle . \tag{104}$$

Fig. 2. Probability density p(u) for the sample coherence $U=Z^2$ for three values of the true coherence: 0.1 (dashed line), 0.5 (solid line), and 0.9 (dot-dashed line). The corresponding true coherences γ^2 are shown. The three frames correspond to different numbers of degrees of freedom.


Degrees of freedom = 5
Degrees of freedom = 10
Degrees of freedom = 50
25
20
15
10
5
0
0.0
0.2
0.4
0.6
0.8
1.0

Based on the previously discussed assumptions, Goodman (1957) found that the probability density function for ϕ is:

$$p(\varphi) = \frac{(1-\gamma^2)^n}{\pi(n-1)!} \sum_{k=0}^{\infty} \frac{2^{k-1}\gamma^k \, \Gamma(n+k/2) \, \Gamma(1+k/2)}{k!} \cos^k\varphi \,, \qquad (105)$$

where φ ranges over the interval $(-\pi, \pi]$. Figure 3 shows $p(\varphi)$ for n = 5, 10 and 50 degrees of freedom, and for γ^2 = 0.1, 0.5 and 0.9.

The joint probability density function for sample coherence and phase is also found in Goodman (1957) (apart from the variable transformation $U = Z^2$):

$$p(u,\varphi) = \frac{(1-\gamma^2)^n}{2\pi(n-1)!(n-2)!} (1-u)^{n-2} \sum_{k=0}^{\infty} \frac{2^k\gamma^k\Gamma^2(n+k/2)}{k!} u^{k/2} \cos^k\varphi. \qquad (106)$$

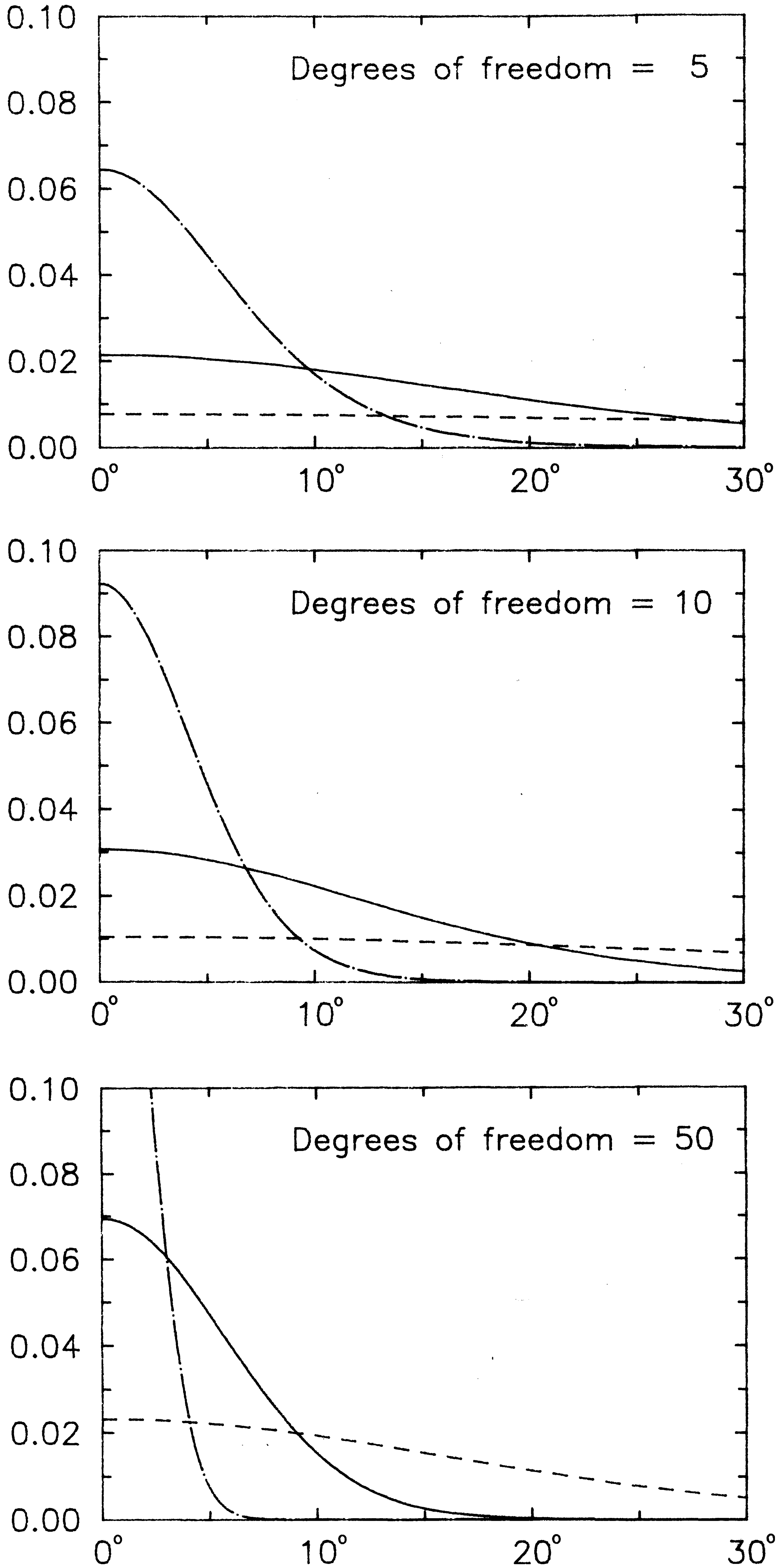
We see from this expression that sample coherence and phase are uncorrelated.

6. COHERENCE AND PHASE STATISTICS

Inspection of Fig. 2 reveals several interesting features of the sample coherence. The probability density function (99) is broader and more skew, the smaller the true coherence or the smaller the degrees of freedom. We suspect that the mean, mode, and median for a particular probability density function characterized by the true coherence γ^2 and degrees of freedom n are all greater than the true coherence, and the more so the smaller γ^2 and n. If this is true there will be a general tendency to overestimate the coherence experimentally, unless this is taken

---

Fig. 3. Probability density function $p(\varphi)$ for the sample phase deviation ϕ for three different values of the true coherence: 0.1 (dashed line), 0.5 (solid line), and 0.9 (dot-dashed line). The three frames correspond to different numbers of degrees of freedom.



Degrees of freedom = 5
Degrees of freedom = 10
Degrees of freedom = 50
0.00
0.02
0.04
0.06
0.08
0.10
0°
10°
20°
30°

into account. Figs. 2 and 3 show that the widths of the probability density functions for the sample coherence (83) and the sample phase (102) are both decreasing functions of γ^2 and n. In this section we will therefore determine all these quantities to develop a practical tool in judging how large n should be in a particular experiment and also possibly in making corrections for bias of the sample-coherence estimate of the true coherence.

First we shall present analytical results for the moments α_1 and α_2, where

$$\alpha_m \equiv \int_0^1 u^m \, p(u) \, du \tag{107}$$

The necessary algebra is somewhat involved, in particular for α_2. For details we refer to Appendix C. Here we just quote the results:

$$\alpha_1 = 1 - \frac{n-1}{n} (1-\gamma^2)^n \, F(n,n;\, n+1;\, \gamma^2) \tag{108}$$

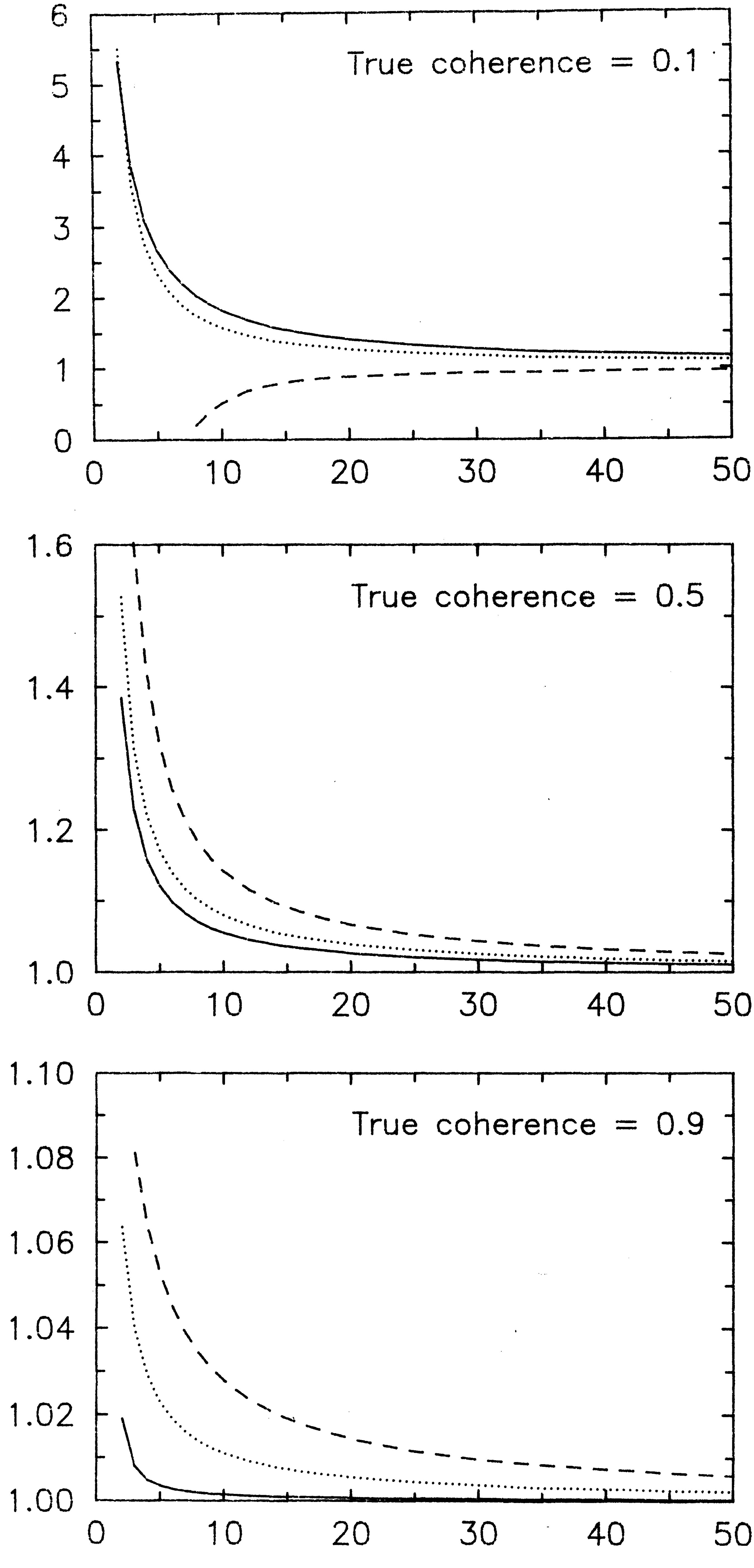
and

$$\alpha_2 = 1 - (1-\gamma^2)^n (n-1) \left\{ \frac{n}{n+1} F(n+1,n;\, n+2;\, \gamma^2) - \frac{n-2}{n} F(n,n;\, n+1;\, \gamma^2) \right\} . \tag{109}$$

Equation (108) is at the same time the expectance of $U = Z^2$,

$$\langle Z^2 \rangle = \alpha_1 \quad , \tag{110}$$

Fig. 4. Ratios of the expectance (solid line), the median (dotted line) and the mode (dashed line) of Z^2 to the true coherence γ^2 as functions of the number of degrees of freedom. The three frames correspond to three different values of γ^2.


True coherence = 0.1
0
1
2
3
4
5
6
0
10
20
30
40
50
True coherence = 0.5
1.0
1.2
1.4
1.6
0
10
20
30
40
50
True coherence = 0.9
1.00
1.02
1.04
1.06
1.08
1.10
0
10
20
30
40
50

whereas the variance of Z^2 is computed from

$$\mathrm{Var}[Z^2] = \alpha_2 - \alpha_1^2 \quad . \tag{111}$$

In devising algorithms for (108) and (109) we found it useful to recast these formulas in terms of the incomplete beta function (see e.g. Kristensen et al. (1983), Appendix A).

$$B_x(a,b) = \int_0^x t^{a-1}(1-t)^{b-1}\,dt \;, \tag{112}$$

for which there is the relation

$$B_x(a,b) = \frac{x^a}{a} F(a,1-b;\; a+1;\; x) \quad . \tag{113}$$

The results were

$$\langle Z^2\rangle = \alpha_1 = 1 - (n-1)\left(\frac{1-\gamma^2}{\gamma^2}\right)^n B_{\gamma^2}(n,1-n) \quad , \tag{114}$$

$$\begin{aligned} \alpha_2 = 1 - (n-1)\left(\frac{1-\gamma^2}{\gamma^2}\right)^n \Big\{ & n\frac{1}{\gamma^2} B_{\gamma^2}(n+1,1-n) \\ & - (n-2)\, B_{\gamma^2}(n,1-n) \Big\} \;, \end{aligned} \tag{115}$$

and

$$\begin{aligned} \mathrm{Var}[Z^2] = (n-1)\Bigg[& \left(\frac{1-\gamma^2}{\gamma^2}\right)^n B_{\gamma^2}(n,1-n) \Big\{ n - \frac{n^2}{\gamma^2} \\ & -(n-1)\left(\frac{1-\gamma^2}{\gamma^2}\right)^n B_{\gamma^2}(n,1-n)\Big\} + n\frac{1-\gamma^2}{\gamma^2} \Bigg] \;. \end{aligned} \tag{116}$$

From (114) and (116) it is possible, after lengthy calculations, sketched in Appendix C, to derive asymptotic results for $\langle Z^2\rangle$ and $\mathrm{Var}[Z^2]$ for large n:

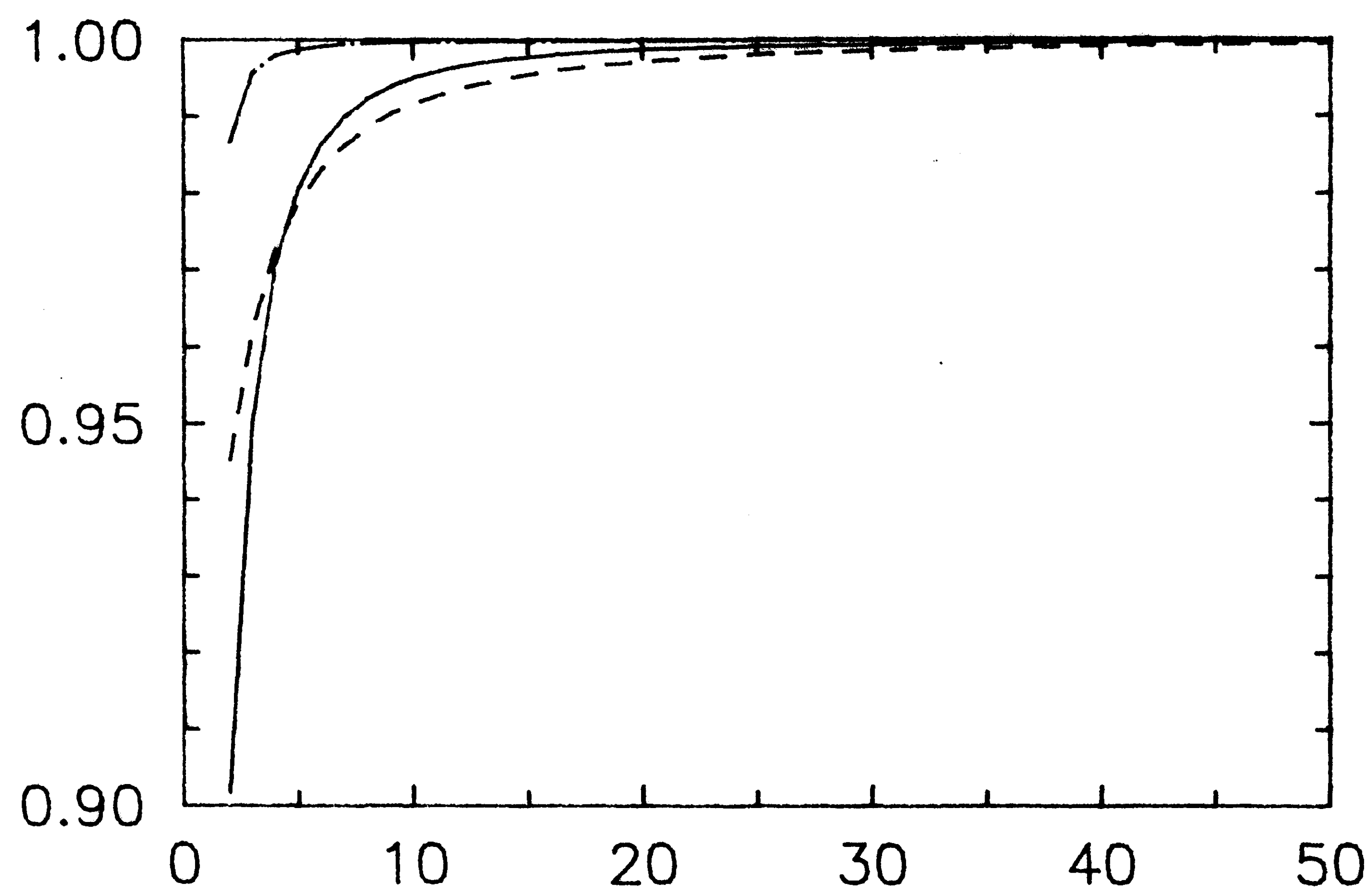


Fig. 5. Ratio of the asymptotic expression (117) to the exact expression (114) as function of the number of degrees of freedom for three values of the true coherence γ^2: 0.1 (dashed line), 0.5 (solid line) and 0.9 (dot-dashed line).

$$\langle Z^2 \rangle = \gamma^2 + \frac{1}{n}(1-\gamma^2)^2 + O(n^{-2}) \qquad (117)$$

and

$$\mathrm{Var}[Z^2] = \frac{1}{n}\, 2\gamma^2(1-\gamma^2)^2 + O(n^{-2}) \quad . \qquad (118)$$

We see that Z^2 is asymptotically unbiased and taken together (117) and (118) express that the limiting form of the probability density function $p(u) = p_n(u)$ is

$$\lim_{n\to\infty} p_n(u) \equiv p_\infty(u) = \delta(u-\gamma^2) \quad , \qquad (119)$$

a result which one would naturally anticipate.

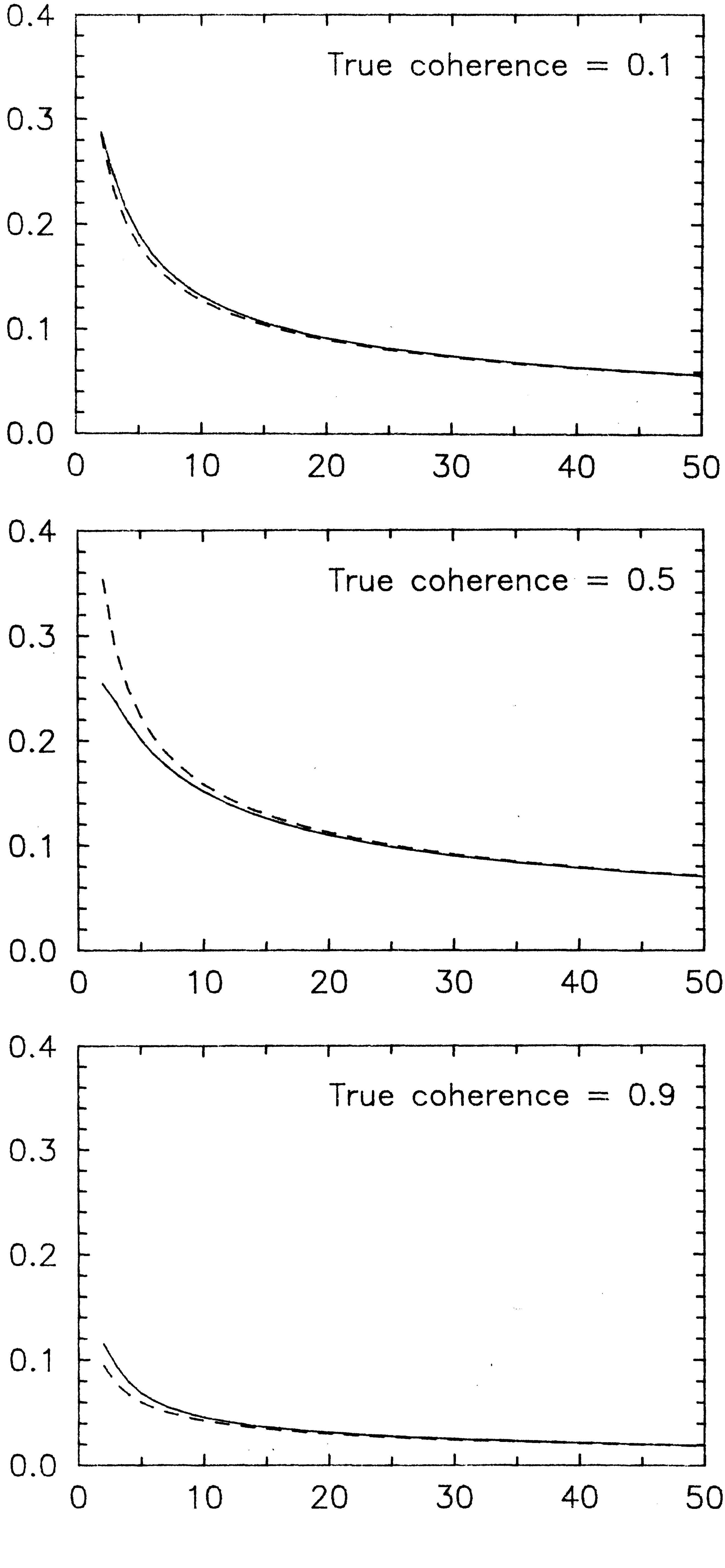
The expressions (114), (116), (117) and (118) are illustrated in Figures 4, 5, and 6. In Fig. 4 we have shown, as functions of the number of degrees of freedom, the expectance $\langle Z^2 \rangle$, the median value of Z^2 (50% fractile) and the mode of Z^2 (the event for which the probability density function has maximum), all divided by the true coherence γ^2. We have selected three different values of γ^2, 0.1, 0.5 and 0.9. From this figure we conclude that, except perhaps for the smallest values of the true coherence, we will overestimate the coherence by just applying "eyeball fitting" to experimental data, no matter whether this fitting is unconsciously based on the expectance, median, or mode. A bias correction to the fit may be obtained by use of (114) or (117). The last equation is considerably easier to use in cases where it is accurate enough. Figure 5 indicates when this is the case. Here we show for γ^2 = 0.1, 0.5, and 0.9 the ratio of expression (117) to (114) as a function of the number of degrees of freedom. Somewhat dependent on the value of γ^2 it seems that if the number of degrees of freedom is greater than about 10 then the approximate equation (117) for $\langle Z^2\rangle$ is sufficiently accurate to estimate the positive bias and the appropriate correction needed.

Figure 6 illustrates (116) and (118). Instead of the variance we have chosen to show the standard deviation, i.e. the square root of $\mathrm{Var}[Z^2]$. Again, we see that for a number of degrees of freedom of 10 or more we can use the asymptotic expression (118) to estimate the standard deviations when we analyse our experiment.

A sketched outline of the deduction of the following results for the phase is given in Appendix C. First we notice that for symmetry reasons we have

$$\langle\phi\rangle = 0 \tag{120}$$

Fig. 6. Standard deviation of Z^2 as function of the number of degrees of freedom for three different values of the true coherence γ^2: 0.1 (top frame), 0.5 (middle frame), and 0.9 (bottom frame). The solid lines correspond to the exact expression (116) and the dashed lines to the asymptotic expression (118).


True coherence = 0.1
True coherence = 0.5
True coherence = 0.9
0.0
0.1
0.2
0.3
0.4
0
10
20
30
40
50

and consequently

$$\mathrm{Var}[\phi] = \langle \phi^2 \rangle \quad . \tag{121}$$

The exact formula for Var $[\phi]$ is complicated:

$$\mathrm{Var}[\phi] = \frac{1}{3}\pi^2 + \frac{2}{(n-1)!}\sum_{k=1}^{\infty}(-1)^k \frac{\Gamma(n+\frac{k}{2})\Gamma(\frac{k}{2})}{(k-1)!\;k^2}\gamma^k F(\frac{k}{2},\frac{k}{2}-n+1;k+1;\gamma^2). \tag{122}$$

In the completely incoherent case $\gamma = 0$ (122) reduces to

$$\mathrm{Var}[\phi] = \frac{1}{3}\pi^2 \tag{123}$$

corresponding to ϕ equidistributed on $(-\pi,\pi]$. For small γ (122) gives the expansion

$$\mathrm{Var}\;[\phi] = \frac{1}{3}\pi^2 - 2\pi^{1/2}\frac{\Gamma(n+\frac{1}{2})}{(n-1)!}\gamma + \frac{n}{2}\gamma^2 + O(\gamma^3) \quad . \tag{124}$$

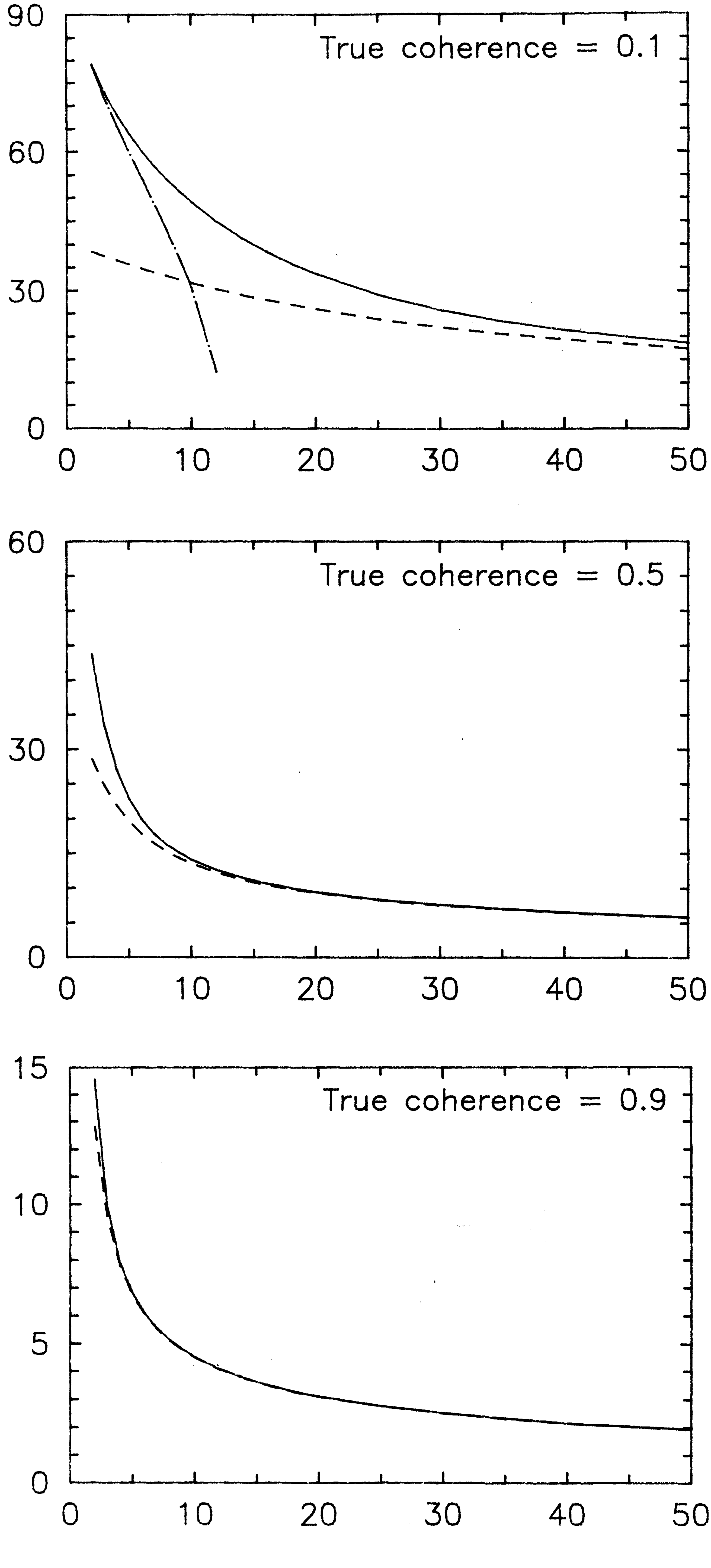
In the perfectly coherent case $\gamma = 1$ it can be shown from (122) that

$$\mathrm{Var}[\phi] = 0 \quad , \tag{125}$$

as was to be expected.

The variance of $\sin\phi$ is simpler to compute than that of ϕ itself. An objection against the use of $\mathrm{Var}[\sin\phi]$ is that in using the sin function we are unable to discern ϕ from its supplement $\pi - \phi$. However, in the most interesting cases we have $\sin\phi \approx \phi$ and the following simple result holds:

Fig. 7. Standard deviation of the phase deviation ϕ (solid line) and $\sin\phi$ (dashed line) as function of the number of degrees of freedom for three different values of the true coherence γ^2: 0.1 (top frame), 0.5 (middle frame), and 0.9 (bottom frame). The dot-dashed line in the top frame shows the square root of the approximate expression (124) for the variance of ϕ.


True coherence = 0.1
90
60
30
0
0
10
20
30
40
50
True coherence = 0.5
60
30
0
0
10
20
30
40
50
True coherence = 0.9
15
10
5
0
0
10
20
30
40
50

$$\mathrm{Var}[\phi] \simeq \mathrm{Var}[\sin\phi] = \frac{1-\gamma^2}{2(n-1)\gamma^2}\,[1-(1-\gamma^2)^{n-1}] \quad . \tag{126}$$

For $\gamma = 1$ we again find that the variance is zero. Equation (126) can also be written as a polynomial in γ^2:

$$\mathrm{Var}[\sin\phi] = \frac{1}{2} + \frac{1}{2(n-1)} \sum_{k=1}^{n-1} (-1)^k \binom{n}{k+1} \gamma^{2k} \quad . \tag{127}$$

For $\gamma^2 \in (0,1)$ and large n we have the asymptotic result

$$\mathrm{Var}[\sin\phi] \simeq \frac{1-\gamma^2}{2n\gamma^2} \quad . \tag{128}$$

For completeness we also quote the statistics for $\cos\phi$:

$$\langle\cos\phi\rangle = \frac{\Gamma(n+1)}{2(n-1)!}\,\pi^{1/2}\,(1-\gamma^2)^n\,\gamma\,F(n+\frac{1}{2},\frac{3}{2};2;\gamma^2) \quad . \tag{129}$$

From (129) and the relation

$$\langle\cos^2\phi\rangle + \mathrm{Var}[\sin\phi] = 1 \tag{130}$$

we may then calculate

$$\mathrm{Var}[\cos\phi] = \langle\cos^2\phi\rangle - \langle\cos\phi\rangle^2 \quad . \tag{131}$$

Figure 7 shows $\mathrm{Var}[\phi]$ and $\mathrm{Var}[\sin\phi]$ as functions of the degrees of freedom for $\gamma^2 = 0.1$, 0.5 and 0.9. In the top frame the approximate expression (124) for $\mathrm{Var}[\phi]$ is also shown. Figure 7 also shows that (124) is a poor approximation except for small values of the number of degrees of freedom and the true coherence.

7. CONCLUSION

We can now return to our original question: How large must the number of degrees of freedom M be to obtain a good estimate of the spectral coherence?

First we note that any smoothing of spectral estimates will in principle destroy information about the fine structure of the spectrum. In other words, the larger we permit M to be in order to improve the statistical confidence the higher will be the chance that we make a simultaneous spectral distortion and lose significant information. As there is an upper limit to the effective number of degrees of freedom M_{eff} for a given record length T, we conclude that there probably is an optimal value of M in most cases for which we can obtain a good statistical estimate without sacrificing too much information. Just how large is this value of M depends on the particular circumstances. We have given a few quantitative tools to aid in this judgment. One is (31), which gives an estimate of the bias that is introduced because of the curvature of the spectrum. This bias is proportional to the square of M. Another is (46), which shows that the standard error of a smoothed spectral estimate is approximately inversely proportional to the square root of M, or rather the square root of M_{eff}. Finally, the effect of the final record length on the statistical dependence between "raw" spectral estimates has been discussed in a refined analysis in section 4. From this we learned that the effective number of degrees of freedom M_{eff} is less than M. We also suggested that an approximate relationship (82) between them exists. Equation (82) shows that the upper limit to M_{eff} is $(T/\mathcal{T})^2/2$, where $\mathcal{T}$ is the integral scale. Often it is possible to obtain an approximate value for $\mathcal{T}$ and we suggest that M is never chosen larger than $(T/\mathcal{T})^2/2$ since the statistical confidence is not improved significantly by using larger values and also there is subsequent penalty in that information is destroyed.

Assuming that we have determined an upper limit to M by taking the considerations above into account, we can use the results from section 6 to make the final decision as to how large M should be. If we specify the standard error, (118) or, if necessary, (116) will tell us, for a given value of γ^2, how large $M = n$ must be. We also showed in section 6 that for a finite value of M we will always overestimate the coherence. It is possible, however, to determine how much this overestimation amounts to by using (117) or, if necessary, (114). These expressions relate the true coherence γ^2 to the sample coherence Z^2 (strictly speaking the ensemble value of Z^2) and they can be solved for γ^2.

If we want to provide confidence intervals we must use the results from section 5, where the probability density functions for the coherence and the phase are given. In this connection it should be pointed out that, in particular for small values of the coherence and the degrees of freedom, the sample coherence is very far from being Gaussian.

With these remarks we consider the discussion about the statistical uncertainties of experimentally determined coherences concluded.

ACKNOWLEDGEMENTS

We are indebted to Dr. Niels Otto Jensen for being a pain in the neck, raising the question about the statistical confidence of sample coherence over and over again. He and our colleague Richard M. Eckman have contributed considerably in the discussions that led to the final formulation of this report. Birthe Skrumsager is acknowledged for her patient typing of the formulas.

REFERENCES

ABRAMOWITZ, M. and STEGUN, I.A. (eds.) (1964). Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables (U.S. Nat. Bur. Stand. Appl. Math. Ser. 55, U.S. Dept. of Commerce, Washington, D.C.).

AMOS, D.E. and KOOPMANS, L.H. (1963). Tables of the Distribution of the Coefficient of Coherence for Stationary Bivariate Gaussian Processes. Sandia Corporation Monograph SCR-483 TID-4500, 328 pp.

BENDAT, J.S. and PIERSOL, A.G. (1971). Random Data: Analysis and Measurement Procedures. (Wiley, New York) xv+407 pp.

COOLEY, J.W. and TUKEY, J.W. (1965). An algorithm for the machine calculation of complex Fourier series. Math. Comput., 19, 297-301.

GOODMAN, N.R. (1957). On the joint estimation of the spectra, cospectrum and quadrature spectrum of a two-dimensional stationary Gaussian process. Sci. paper 10, Engineering Statistics Laboratory, New York University.

GOODMAN, N.R. (1963). Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction). Ann. Math. Statist., 34, 152-177.

GRADSHTEYN, I.S. and RYZHIK, I.M. (1980). Tables of Integrals, Series and Products. English version ed. by Alan Jeffrey, (Academic Press, New York) 1160 pp.

KOOPMANS, L.H. (1974). The Spectral Analysis of Time Series. (Academic Press, New York) xiv+368 pp.

KRISTENSEN, L., KIRKEGAARD, P. and LENSCHOW, D.H. (1983). Squashed Atmospheric Turbulence. Risø Report No. 478, 79 pp.

LUMLEY, J.L. and PANOFSKY, H.A. (1964). The Structure of Atmospheric Turbulence. Interscience Monographs and Texts in Physics and Astronomy (Wiley, New York) xi+239 pp.

PANOFSKY, H.A. and DUTTON, J.A. (1984). Atmospheric Turbulence. Models and Methods for Engineering Applications. Wiley-Interscience Publications (Wiley, New York) xix+397 pp.

APPENDIX

We collect here a number of mathematical auxiliaries which with their rather technical nature would be inappropriate in the main text. The following material, which falls naturally into three parts, may serve to document our findings; some of the equations and their derivation might be interesting in their own right.

A. Trigonometric Integrals

We shall prove the following two integral formulas:

$$\int_{-\infty}^{\infty} \frac{\sin(x'-x)}{x'-x}\,\frac{\sin(x''-x)}{x''-x}\,dx = \pi\,\frac{\sin(x''-x')}{x''-x'} \tag{A1}$$

and

$$\int_{-\infty}^{\infty} \frac{\sin(x'-x)}{x'-x}\,\frac{\sin(x''-x)}{x''-x}\,\frac{1}{1+(x/\theta)^2}\,dx =$$

$$\frac{\pi\theta}{2(\theta^2+x'^2)(\theta^2+x''^2)}\left\{\theta(2\theta^2+x'^2+x''^2)\,\frac{\sin(x''-x')}{x''-x'}\right.$$

$$-\ (\theta^2-x'x'')\cos(x''-x')$$

$$\left. +\ e^{-2\theta}\left[(\theta^2-x'x'')\cos(x'+x'') - \theta(x'+x'')\sin(x'+x'')\right]\right\}. \tag{A2}$$

In (A1) and (A2) x' and x'' are real numbers, and θ is real and positive. Not even (A1) could be found in standard integral tables. Of course (A1) comes out as a limiting case of (A2) when $\theta \to \infty$, once the latter is established. But (A1) is easily proved directly: The first factor in the integrand can be written

$$\frac{\sin(x'-x)}{x'-x} = \frac{1}{2}\int_{-1}^{1} e^{i(x'-x)t}\,dt\ , \tag{A3}$$

and similarly for the other. By reversing the order of integration, introduction of δ functions, and once more use of (A3), (A1) follows. The proof of (A2) is more complicated. We write the left-hand side of J of (A2) as the sum of two Cauchy principal-value integrals

$$J = J_1 + J_2 \tag{A4}$$

where

$$J_1 = \frac{1}{2} \cos 2a \; ⨍_{-\infty}^{\infty} \frac{1}{1 + (\frac{y+b}{\theta})^2} \frac{1}{y^2 - a^2} dy \tag{A5}$$

and

$$J_2 = -\frac{1}{2} ⨍_{-\infty}^{\infty} \frac{1}{1 + (\frac{y+b}{\theta})^2} \frac{\cos 2y}{y^2 - a^2} dy \quad , \tag{A6}$$

and where we have introduced the new parameters

$$a = \frac{1}{2} (x'' - x') \tag{A7}$$

and

$$b = \frac{1}{2} (x' + x'') \quad . \tag{A8}$$

To evaluate (A5) and (A6) let us consider the integral

$$J_3 = J_3(\alpha, \beta, c) = \int_{-\infty}^{\infty} Q(x) \cos x \, dx, \tag{A9}$$

where

$$Q(z) = \frac{1}{(z-\alpha)^2 + \beta^2} \frac{1}{z^2 + c^2} \quad . \tag{A10}$$

α and β are real with $\beta > 0$; we temporarily assume c real and positive. $Q(z)e^{iz}$ is analytic in the halfplane $\mathrm{Im}(z) > 0$ except for simple poles at $z = ic$ and $z = \alpha + i\beta$ where the residues are

$$R_1 = \frac{1}{(ic-\alpha)^2 + \beta^2} \frac{e^{-c}}{2ic} \tag{A11}$$

and

$$R_2 = \frac{1}{2i\beta} \frac{e^{-\beta+i\alpha}}{(\alpha+i\beta)^2 + c^2} \quad . \tag{A12}$$

Contour integration yields

$$\int_{-\infty}^{\infty} Q(x)\, e^{ix}\, dx = 2\pi i(R_1 + R_2) \quad , \tag{A13}$$

and if we take the real part of (A13) we get

$$J_3 = \pi \frac{\beta e^{-c}(\alpha^2+\beta^2-c^2)+ce^{-\beta}[(\alpha^2-\beta^2+c^2)\cos\alpha+2\alpha\beta\sin\alpha]}{\beta c[(\alpha^2+\beta^2+c^2)^2 - 4\beta^2c^2]} \quad , \tag{A14}$$

which, by analytical continuation, extends to all complex c for which (A14) is finite. We now apply the operator $J_3(\alpha,\beta,i\gamma)$ + $J_3(\alpha,\beta,-i\gamma)$ to (A14) and substitute $k(\alpha,\beta,\gamma,x)$ for (α,β,γ,x) in the integral (A9) ($k > 0$). Then we get

$$⨍_{-\infty}^{\infty} \frac{1}{(x-\alpha)^2+\beta^2} \frac{\cos kx}{x^2-\gamma^2}\, dx =$$

$$\pi \frac{-\beta\sin k\gamma(\alpha^2+\beta^2+\gamma^2) + \gamma e^{-k\beta}[(\alpha^2-\beta^2-\gamma^2)\cos k\alpha+2\alpha\beta\sin k\alpha]}{\beta\gamma[(\alpha^2+\beta^2-\gamma^2)^2 + 4\beta^2\gamma^2]} \tag{A15}$$

In particular we find for $k \to 0$:

$$⨍_{-\infty}^{\infty} \frac{1}{(x-\alpha)^2+\beta^2} \frac{1}{x^2-\gamma^2}\, dx = \frac{\pi}{\beta} \frac{\alpha^2-\beta^2-\gamma^2}{(\alpha^2+\beta^2-\gamma^2)^2+4\beta^2\gamma^2} \quad . \tag{A16}$$

(A16) and (A15) are applied to express (A5) and (A6):

$$J_1 = \frac{1}{2} \pi\theta \frac{b^2-\theta^2-a^2}{(b^2+\theta^2-a^2)^2 + 4\theta^2a^2} \cos 2a \tag{A17}$$

and

$$J_2 = \frac{\pi\theta}{2a} \frac{\theta\sin 2a(b^2+\theta^2+a^2) - ae^{-2\theta}\left[(b^2-\theta^2-a^2)\cos 2b+2b\theta\sin 2b\right]}{(b^2+\theta^2-a^2)^2 + 4\theta^2a^2}. \tag{A18}$$

Going back to the original parameters x' and x" we arrive at (A2). Numerical check calculations confirm the correctness of this expression.

B. Reduction of Degrees of Freedom. Alternative Approach

In section 4 it was mentioned that the spectrum (79) also could be analysed to yield information about the error variance and effective number of degrees of freedom. The method we shall use may be seen as an interesting alternative to the outline in section 4 for the Cauchy case. It involves a temporary detour to the time domain, exploiting the Fourier-transform duality between power spectra and autocovariance functions:

$$\phi(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} R(\tau)\, e^{-i\omega\tau}\, d\tau \quad , \tag{B1}$$

$$R(\tau) = \int_{-\infty}^{\infty} \phi(\omega)\, e^{i\omega\tau}\, d\omega \quad . \tag{B2}$$

We begin with (39) with x = y:

$$E_{xx}[k',k";N] = E[k',k";N] \simeq$$

$$(-1)^{k"-k'}\, e^{i\pi \frac{k"-k'}{N}} \int_{-\infty}^{\infty} \phi(\omega) \frac{\sin(\pi k'-\omega T/2)}{\pi k'-\omega T/2} \frac{\sin(\pi k"-\omega T/2)}{\pi k"-\omega T/2}\, d\omega. \tag{B3}$$

Using (B1) and (A3) we obtain

$$E[k',k'';N] =$$

$$(-1)^{k''-k'}\, e^{i\pi\frac{k''-k'}{N}}\, \frac{1}{T^2} \quad \times$$

$$\int_{-\frac{T}{2}}^{\frac{T}{2}} dt' \int_{-\frac{T}{2}}^{\frac{T}{2}} dt'' R(t''-t')\, e^{-i\frac{2\pi}{T}(k''t''-k't')} \qquad \text{(B4)}$$

This integral is essentially the same as (4) and hence can be recast immediately in the form (9):

$$E[k',k'';N] = (-1)^{k''-k'}\, e^{i\pi\frac{k''-k'}{N}}\, \frac{1}{T} \quad \times$$

$$\int_{-T}^{T} R(\tau)\, \frac{\sin(\pi(k''-k')(1-\frac{|\tau|}{T}))}{\pi(k''-k')}\, e^{-i\pi(k'+k'')\frac{\tau}{T}} \quad d\tau\ . \qquad \text{(B5)}$$

$$= \delta_{k'k''}\, \frac{1}{T} \int_{-T}^{T} R(\tau) e^{-i\pi(k'+k'')\frac{\tau}{T}}\, d\tau$$

$$- \frac{e^{i\pi\frac{k''-k'}{N}}}{\pi(k''-k')}\, \frac{1}{T} \int_{-T}^{T} R(\tau)\, \sin(\pi(k''-k')\frac{|\tau|}{T})\, e^{-i\pi(k'+k'')\frac{\tau}{T}}\, d\tau. \qquad \text{(B6)}$$

If the integral scale $\mathcal{T} << T$ we have

$$E[k',k'';N] \simeq$$

$$\Delta\omega \left\{ \phi(\frac{\pi}{T}(k'+k''))\, \delta_{k'k''} - \frac{e^{i\pi\frac{k''-k'}{N}}}{\pi(k''-k')} \right.$$

$$\times \frac{1}{2\pi} \int_{-\infty}^{\infty} R(\tau)\, \sin(\pi(k''-k')\, \frac{|\tau|}{T})\, e^{-i\pi(k'+k'')\frac{\tau}{T}}\, d\tau. \qquad \text{(B7)}$$

We see from (35) that we need to evaluate $|E[k',k'';N]|^2$ and $|E[k',-k'';N]|^2$. Thus we get

$$|E[k',k'';N]|^2 = (\Delta\omega)^2 \Big\{ \phi(\frac{2\pi}{T} k') \Big[\phi(\frac{2\pi}{T}k') -$$

$$\frac{2}{\pi} \int_0^\infty R(\tau) \frac{\tau}{T} \cos(\frac{2\pi}{T} k'\tau)\, d\tau \Big] \delta_{k'k''}$$

$$+ \frac{1}{\pi^2(k''-k')^2}\Big[\frac{1}{2\pi} \int_{-\infty}^{\infty} R(\tau)\sin(\pi(k''-k')\frac{|\tau|}{T}) \cos(\pi(k'+k'')\frac{\tau}{T})\, d\tau\Big]^2 \Big\} \tag{B8}$$

and

$$|E[k',-k'';N]|^2 = (\Delta\omega)^2 \frac{1}{\pi^2(k'+k'')^2} \times$$

$$\Big[\frac{1}{2\pi} \int_{-\infty}^{\infty} R(\tau)\sin(\pi(k'+k'')\frac{|\tau|}{T}) \cos(\pi(k''-k')\frac{\tau}{T})\, d\tau\Big]^2 \quad . \tag{B9}$$

Note that there is only one term in (B9) since according to (35) both k' and k" are positive. We need the sum S of (B8) and (B9) and write it in the form

$$S[k',k'';N] \equiv |E[k',k'';N]|^2 + |E[k',-k'';N]|^2 =$$

$$(\Delta\omega)^2 \Big\{\phi(\frac{2\pi}{T}k') \Big[\phi(\frac{2\pi}{T}k') - \frac{2}{\pi} \int_0^\infty R(\tau) \frac{\tau}{T} \cos(\frac{2\pi}{T}k'\tau)\, d\tau\Big] \delta_{k'k''}$$

$$+\frac{1}{\pi^2} \frac{1}{(k''-k')^2}\Big[\frac{1}{2\pi} \int_0^\infty R(\tau)\sin(\frac{2\pi}{T}k''\tau)d\tau - \frac{1}{2\pi} \int_0^\infty R(\tau)\sin(\frac{2\pi}{T}k'\tau)d\tau \Big]^2$$

$$+\frac{1}{\pi^2} \frac{1}{(k'+k'')^2}\Big[\frac{1}{2\pi} \int_0^\infty R(\tau)\sin(\frac{2\pi}{T}k'\tau)d\tau + \frac{1}{2\pi} \int_0^\infty R(\tau)\sin(\frac{2\pi}{T}k''\tau)d\tau \Big]^2\Big\}. \tag{B10}$$

The integrals in (B10) are cosine and sine-transforms of $R(\tau)\tau/T$ and $R(\tau)$, respectively; the first follows from the second by derivation, but neither are sufficient to reconstruct $R(\tau)$ and $\phi(\omega)$.

In view of this we are unable to pursue the general development further and are forced to assume, as in section 4, a particular form of the spectrum in order to obtain a useful expression for the error variance. We take (79) - (80) as a slightly more general spectral shape than (55). Note that $\alpha = 5/6$ corresponds to the von Karman spectrum and $\alpha = 1$ to the Cauchy case. Using (B2) we get

$$R(\tau) = \frac{2\sigma^2}{\Gamma(\alpha-\frac{1}{2})} \left(\frac{a}{2}\frac{\tau}{\mathcal{T}}\right)^{\alpha-\frac{1}{2}} K_{\alpha-\frac{1}{2}}\left(a\frac{\tau}{\mathcal{T}}\right) , \tag{B11}$$

where K_ν is the modified Bessel function of the second kind (for convenience we assume $\tau > 0$). If we substitute (B11) in (B10) and use the two formulas (see e.g. Gradshteyn and Ryzhik, 1980, p. 747):

$$\int_0^\infty x^{\alpha+\frac{1}{2}} K_{\alpha-\frac{1}{2}}(x)\cos\beta x\,dx = 2^{\alpha-\frac{1}{2}}\Gamma(\alpha+\tfrac{1}{2})F(\alpha+\frac{1}{2},1;\frac{1}{2};-\beta^2) \tag{B12}$$

and

$$\int_0^\infty x^{\alpha-\frac{1}{2}} K_{\alpha-\frac{1}{2}}(x)\sin\beta x\,dx = 2^{\alpha-\frac{1}{2}}\beta\Gamma(\alpha+\tfrac{1}{2})F(\alpha+\frac{1}{2},1;\frac{3}{2};-\beta^2) \tag{B13}$$

we find

$S[k',k";N] =$

$$(\Delta\omega)^2 \Bigg\{\phi\left(\frac{2\pi}{T}k'\right)\left[\phi\left(\frac{2\pi}{T}k'\right) - \frac{1}{\theta}\,\frac{\sigma^2\mathcal{T}}{\pi}\,\frac{2\alpha-1}{a^2}\,F(\alpha+\frac{1}{2},1,\frac{1}{2};-\beta'^2)\right]\delta_{k'k"}$$

$$+\left[\frac{1}{2\theta}\,\frac{\sigma^2\mathcal{T}}{\pi}\,\frac{2\alpha-1}{a^2}\right]^2\Bigg[\left\{\frac{\beta"F(\alpha+\frac{1}{2},1;\frac{3}{2};-\beta"^2)-\beta'F(\alpha+\frac{1}{2},1,\frac{3}{2};-\beta'^2)}{\beta"-\beta'}\right\}^2$$

$$+\left\{\frac{\beta'F(\alpha+\frac{1}{2},1;\frac{3}{2};-\beta'^2)+\beta"F(\alpha+\frac{1}{2},1;\frac{3}{2};-\beta"^2)}{\beta'+\beta"}\right\}^2\Bigg]\Bigg\} , \tag{B14}$$

where

$$\left\{ \begin{matrix} \beta' \\ \beta'' \end{matrix} \right\} = \frac{2\pi}{T} \frac{\mathcal{T}}{a} \left\{ \begin{matrix} k' \\ k'' \end{matrix} \right\} \tag{B15}$$

and

$$\phi(\frac{2\pi}{T}k') = \frac{\sigma^2\mathcal{T}}{\pi} \frac{1}{(1+\beta'^2)^\alpha} \quad . \tag{B16}$$

We first evaluate $S[k',k'';N]$ for small and large values of β' and β''.

<u>$\beta' << 1$ and $\beta'' << 1$:</u>

In this limit we have

$$F(\alpha + 1/2, 1; n/2; -\beta^2) \simeq 1 \tag{B17}$$

so that

$$S[k',k'';N] \simeq (\Delta\omega)^2 \left(\frac{\sigma^2\mathcal{T}}{\pi}\right)^2 \left\{\left(1 - \frac{2\alpha-1}{a^2} \frac{1}{\theta}\right) \delta_{k'k''} + \frac{1}{2} \left(\frac{2\alpha-1}{a^2} \frac{1}{\theta}\right)^2\right\} \tag{B18}$$

<u>$\beta' >> 1$ and $\beta'' >> 1$:</u>

Here we use the well-known continuation formula for hypergeometric functions:

$$F(a,b;c;z) = \frac{\Gamma(c)\Gamma(b-a)}{\Gamma(b)\Gamma(c-a)} (-z)^{-a} F(a, 1-c+a; 1-b+a; z^{-1})$$
$$+ \frac{\Gamma(c)\Gamma(a-b)}{\Gamma(a)\Gamma(c-b)} (-z)^{-b} F(b, 1-c+b; 1-a+b; z^{-1}) \tag{B19}$$

This gives

$$F(\alpha+\frac{1}{2},1,\frac{n}{2};-\beta^2) =$$

$$(1+O(\beta^{-2}))\left\{\frac{\Gamma(\frac{n}{2})\Gamma(\frac{1}{2}-\alpha)}{\Gamma(\frac{n-1}{2}-\alpha)}\beta^{-2\alpha-1} + \frac{n-2}{2\alpha-1}\beta^{-2}\right\}. \quad (n=1,3) \tag{B20}$$

We shall assume $\alpha > 1/2$, and keeping only the dominant term we get

$$F(\alpha+\frac{1}{2},1;\frac{n}{2};-\beta^2) \approx \frac{n-2}{2\alpha-1}\beta^{-2} \quad . \tag{B21}$$

Further, we have for large β

$$\phi(\omega) \approx \frac{\sigma^2}{\pi}\beta^{-2\alpha} \tag{B22}$$

Substituting in (B14) we get

$$S[k',k'';N] \approx (\Delta\omega)^2\left(\frac{\sigma^2}{\pi}\right)^2\left\{\beta'^{-4\alpha}\left[1+\frac{\beta'^{2(\alpha-1)}}{a^2\theta}\right]\delta_{k'k''} + \frac{1}{2}\frac{1}{(a^2\theta)^2}\frac{1}{\beta'^2}\frac{1}{\beta''^2}\right\} \tag{B23}$$

Combining (B18) and (B23) and reintroducing (B16) we can write

$$S[k',k'';N] = \left[\phi(\frac{2\pi}{T}k')\Delta\omega\right]^2 \times$$

$$\begin{cases} \left(1 - \frac{2\alpha-1}{a^2\theta}\right)\delta_{k'k''} + \frac{1}{2}\left(\frac{2\alpha-1}{a^2\theta}\right)^2, & \beta',\beta'' \ll 1 \\ \left(1 + \frac{\beta'^{2(\alpha-1)}}{a^2\theta}\right)\delta_{k'k''} + \frac{1}{2}\frac{\beta'^{2(\alpha-1)}\beta''^{2(\alpha-1)}}{(a^2\theta)^2}, & \beta',\beta'' \gg 1 \end{cases} \tag{B24}$$

The error variance (35) becomes

$$\sigma^2[k;N;M] = \frac{1}{M^2} \sum_{m'=0}^{M-1} \sum_{m''=0}^{M-1} S[k+m',k+m'';N] \simeq \left\{ \phi(\omega_*)\Delta\omega \right\} \times$$

$$\begin{cases} (1-\dfrac{2\alpha-1}{a^2\theta})\dfrac{1}{M} + \dfrac{1}{2}(\dfrac{2\alpha-1}{a^2\theta})^2 , & k \ll \dfrac{a}{2\pi}\dfrac{T}{\mathcal{T}} \\ (1+\dfrac{1}{a^2\theta}(\dfrac{\omega_*\mathcal{T}}{a})^{2(\alpha-1)})\dfrac{1}{M} + \dfrac{1}{2}(\dfrac{\omega_*\mathcal{T}}{a})^{4(\alpha-1)}\dfrac{1}{a^4\theta^2} , & k \gg \dfrac{a}{2\pi}\dfrac{T}{\mathcal{T}}, \end{cases}$$

(B25)

where ω_* is given by (74). We assume, as stated before, that $\theta \gg 1$, but not necessarily that $\theta \gg \sqrt{M}$. Therefore, for $a \leqslant 1$, we can write

$$\sigma^2[k;N;M] \simeq \left\{ \phi(\omega_*)\Delta\omega \right\}^2 \cdot$$

$$\begin{cases} \dfrac{1}{M} + 2(\dfrac{2\alpha-1}{a^2})^2 (\dfrac{\mathcal{T}}{T})^2 & \text{for } \dfrac{\omega_*\mathcal{T}}{a} \ll 1 \\ \dfrac{1}{M} + 2\dfrac{1}{a^4}(\dfrac{\omega_*\mathcal{T}}{a})^{4(\alpha-1)} (\dfrac{\mathcal{T}}{T})^2 & \text{for } \dfrac{\omega_*\mathcal{T}}{a} \gg 1 \end{cases} \qquad \text{(B26)}$$

We see that when α is greater than 1/2 and also bounded away from this limit then (78) is a conservative estimate in both frequency limits.

C. Derivation of Statistics in Goodman Distributions

In the following we shall give the necessary justifications of the results presented in section 6.

Equation (108) is derived from (99) and (107) by term-by-term integration, giving

$$\alpha_1 = \frac{1}{(n-1)!} (1-\gamma^2)^n \sum_{k=0}^{\infty} \frac{(k+n-1)!}{k!} \frac{k+1}{k+n} \gamma^{2k} , \tag{C1}$$

followed by use of the identity

$$\frac{k+1}{k+n} = 1 - \frac{n-1}{k+n} . \tag{C2}$$

Similarly we find

$$\alpha_2 = \frac{1}{(n-1)!} (1-\gamma^2)^n \sum_{k=0}^{\infty} \frac{(k+n-1)!}{k!} \frac{(k+1)(k+2)}{(k+n)(k+n+1)} \gamma^{2k}, \tag{C3}$$

which by the identity

$$\frac{(k+1)(k+2)}{(k+n)(k+n+1)} = 1 - \frac{n-1}{k+n+1} - \frac{n-1}{k+n} + \frac{(n-1)^2}{(k+n)(k+n+1)} \tag{C4}$$

is transformed into

$$\alpha_2 = 1 - (1-\gamma^2)^n \left\{ \frac{n-1}{n+1} F(n+1,n;n+2;\gamma^2) + \frac{n-1}{n} F(n,n;n+1;\gamma^2) - \frac{(n-1)^2}{n(n+1)} F(n,n;n+2;\gamma^2) \right\} . \tag{C5}$$

The last F function can be expressed in terms of the two others by aid of the fifth Gaussian contiguity relation:

$$(c-a-1)F + aF(a+1) - (c-1)F(c-1) = 0, \tag{C6}$$

resulting in (109).

From (111), (114) and (115), we compute

$$\mathrm{Var}[Z^2] = (n-1)\left(\frac{1-\gamma^2}{\gamma^2}\right)^n \left[B_{\gamma^2}(n,1-n)\left\{n-(n-1)\left(\frac{1-\gamma^2}{\gamma^2}\right)^n B_{\gamma^2}(n,1-n)\right\} - \frac{n}{\gamma^2} B_{\gamma^2}(n+1,1-n) \right] , \tag{C7}$$

and equation (116) is obtained from (C7) by use of the recurrence relation

$$B_x(a+1,b) = \frac{a}{a+b} B_x(a,b) - \frac{1}{a+b} x^a(1-x)^b \quad , \tag{C8}$$

which is derived from Abramowitz and Stegun's (1964) formula 26.5.16 p.944; in this way $B_{\gamma^2}(n+1,1-n)$ is eliminated.

In the derivation of the asymtotic expressions (117) and (118) let us write x for γ^2 for convenience and introduce the quantity

$$U(n,x) = (n-1)\left(\frac{1-x}{x}\right)^n B_x(n,1-n)$$

$$= \frac{n-1}{n} (1-x)^n F(n,n;n+1;x) \tag{C9}$$

By Kummer's relation,

$$F(a,b;c;x) = (1-x)^{-a} F(a,c-b;c;\frac{x}{x-1}) \tag{C10}$$

(C9) is transformed to

$$U(n,x) = \frac{n-1}{n} F(n,1;n+1;\frac{x}{x-1}) \quad . \tag{C11}$$

Now, if $x \in [0,1/2)$(this restriction is relieved in the final results by analytical continuation), (C11) can be expressed as the convergent series

$$U(n,x) = (n-1) \sum_{k=0}^{\infty} \frac{1}{n+k} \left(\frac{x}{x-1}\right)^k \quad . \tag{C12}$$

We shall make use of the expansion

$$\frac{1}{n+k} = \frac{1}{n} \left[\sum_{r=0}^{s} (-1)^r \frac{k^r}{n^r} + O\left(\frac{1}{n^{s+1}}\right)\right] \tag{C13}$$

and the identities

$$\sum_{k=0}^{\infty} y^k = (1-y)^{-1} \quad , \tag{C14}$$

$$\sum_{k=0}^{\infty} ky^k = y(1-y)^{-2} \quad , \tag{C15}$$

$$\sum_{k=0}^{\infty} k^2y^k = y(1+y)(1-y)^{-3} \quad , \tag{C16}$$

$$\sum_{k=0}^{\infty} k^3y^k = y(1-y)^{-4} (1+4y+y^2) \quad , \tag{C17}$$

all of which can be deduced from the summation formula

$$\sum_{k=0}^{\infty} k(k-1)\ldots(k-r+1)\, y^k = r!y^r(1-y)^{-r-1} \quad . \tag{C18}$$

To evaluate $\langle Z^2 \rangle$ as given in (114) we let s = 1 in (C13), such that (C12) gives

$$U = U(n,x) = \frac{n-1}{n} \sum_{k=0}^{\infty} (1-\frac{k}{n})\, y^k + O(n^{-2}), \tag{C19}$$

where

$$\left.\begin{aligned} y &= \frac{x}{x-1} \qquad &(C20a) \\ x &= \frac{y}{y-1} \qquad &(C20b) \end{aligned}\right\}$$

By (C14) and (C15), (C19) becomes

$$U = (1-y)^{-1} - \frac{1}{n}(1-y)^{-2} + O(n^{-2}) \quad . \tag{C21}$$

We get

$$\langle Z^2\rangle = 1-U = x + \frac{1}{n}(1-x)^2 + O(n^{-2}) \ , \tag{C22}$$

which is equivalent to (117). To derive (118) it is necessary to take s = 3 in (C13). Then

$$U = \frac{n-1}{n}\sum_{k=0}^{\infty}\left(1 - \frac{k}{n} + \frac{k^2}{n^2} - \frac{k^3}{n^3}\right) y^k + O(n^{-4}) \ , \tag{C23}$$

and by (C14) - (C17), we obtain

$$U = (1-y)^{-1} - \frac{1}{n}(1-y)^{-2} + \frac{1}{n^2}2y(1-y)^{-3}$$

$$- \frac{1}{n^3}2y(1-y)^{-4}(1+2y) + O(n^{-4}) \quad . \tag{C24}$$

From (116), (C9) and (C20) we get

$$\mathrm{Var}[Z^2] = U\left(\frac{1-y}{y}n^2 + n - U\right) - n^2/y + n/y \ . \tag{C25}$$

When (C24) is inserted in (C25) we see, after reduction, that

$$\mathrm{Var}[Z^2] = -\frac{1}{n}2y(1-y)^{-3} + O(n^{-2}) \quad , \tag{C26}$$

which is equivalent to (118).

Next we shall consider the calculations involving the phase statistics. We first observe that there is an alternative expression in Goodman (1957) to $p(\varphi)$ in (105):

$$p(\varphi) = \frac{(1-\gamma^2)^n}{2\pi}\left\{1 - \frac{ns}{(1-s^2)^{n+\frac{1}{2}}}\left[\frac{\Gamma(\frac{1}{2})\Gamma(n+\frac{1}{2})}{n!} \pm B_{s^2}(\tfrac{1}{2},n+\tfrac{1}{2})\right]\right\}, \tag{C27}$$

where

$$s = -\gamma\cos\varphi \tag{C28}$$

In (C27) the sign "+" should be selected if $|\varphi| \in [0,\frac{1}{2}\pi]$, and "-" if $|\varphi| \in [\frac{1}{2}\pi,\pi]$. Equations (105) and (C27) were derived by contraction of joint probability densities over different coordinate variables. It is possible to resolve the sign ambiguity in (C27) by (113):

$$B_{s^2}(1/2,n+1/2) = 2|s|\ F(1/2,1/2-n;3/2;s^2) \quad . \tag{C29}$$

This results in

$$p(\varphi) = \frac{(1-\gamma^2)^n}{2\pi}\left\{1 - \frac{ns}{(1-s)^{n+\frac{1}{2}}}\left[\frac{\Gamma(\frac{1}{2})\Gamma(n+\frac{1}{2})}{n!} - 2sF(1/2,1/2-n;3/2;s^2)\right]\right\} . \tag{C30}$$

Equation (C27) is well suited for a calculation of p( ) itself, but we prefer (105) for moments calculations. For the variance of ϕ we find

$$\mathrm{Var}[\phi] = \frac{(1-\gamma^2)^n}{\pi(n-1)!}\sum_{k=0}^{\infty}\frac{2^{k-1}\gamma^k\ \Gamma(n+k/2)\Gamma(1+k/2)}{k!} I_k, \tag{C31}$$

where

$$I_k \equiv \int_{-\pi}^{\pi} \varphi^2 \cos^k\varphi \, d\varphi = (-1)^k \frac{2\pi}{2^k} \left\{ \sum_{r=0}^{[\frac{k-1}{2}]} \binom{k}{r} \frac{1}{(\frac{k}{2}-r)^2} + a_k \right\} \tag{C32}$$

$$a_k = \begin{cases} \binom{k}{\frac{k}{2}} \frac{\pi^2}{3} , & k \text{ even} \\ 0, & k \text{ odd} \end{cases}$$

Equation (C32) follows from the expansion

$$\cos^k \varphi = \left(\frac{e^{i\varphi}+e^{-i\varphi}}{2}\right)^k = 2^{-k} \sum_{r=0}^{k} \binom{k}{r} e^{(k-2r)i\varphi} \tag{C33}$$

in conjunction with

$$\int_{-\pi}^{\pi} \varphi^2 e^{is\varphi} \, d\varphi = \int_{-\pi}^{\pi} \varphi^2 \cos s\varphi \, d\varphi = \begin{cases} 2\pi^2/3 & s = 0 \\ (-1)^s \frac{4\pi}{s^2} & s \neq 0 \end{cases} \tag{C34}$$

When (C32) is inserted in (C31) we arrive, after some algebra, which includes the identity

$$\frac{1}{(n-1)!} \sum_{m=0}^{\infty} \frac{(n+m-1)!}{m!} \gamma^{2m} = (1-\gamma^2)^{-n} \quad , \tag{C35}$$

to the expression

$$\mathrm{Var}[\phi] =$$

$$\frac{1}{3}\pi^2 + \frac{2(1-\gamma^2)^n}{(n-1)!} \sum_{k=1}^{\infty} (-1)^k \frac{\Gamma(n+\frac{k}{2})\Gamma(\frac{k}{2})}{(k-1)! \; k^2} \gamma^k F(n+\frac{k}{2}, 1+\frac{k}{2}, k+1; \gamma^2), \tag{C36}$$

which in turn is transformed to (122), when (100) is used. To derive (125) (the perfectly coherent case) we use the well-known identities

$$F(a,b;c;1) = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)} \tag{C37}$$

and

$$\sum_{k=1}^{\infty} (-1)^k \frac{1}{k^2} = -\frac{\pi^2}{12} \tag{C38}$$

The variance of $\sin\phi$ is computed in the same way as $\mathrm{Var}[\phi]$, but the algebra is simpler and leads to the series

$$\mathrm{Var}[\sin\phi] = \frac{1}{2} \frac{(1-\gamma^2)^n}{(n-1)!} \sum_{m=0}^{\infty} \frac{(n+m-1)!}{(m+1)!} \gamma^{2m} \quad . \tag{C39}$$

This can be reduced further, as we infer from (C35)

$$\sum_{m=0}^{\infty} \frac{(n+m-1)!}{(m+1)!} \gamma^{2m} = (n-2)! \frac{1}{\gamma^2} \left[(1-\gamma^2)^{-n+1} - 1\right] \tag{C40}$$

whereafter (C39) can be written in the form (126).

We shall finally give a short account of the numerical procedures used for evaluating the three location parameters for the sample coherence Z^2: expectance, median and mode. These were mentioned in section 6 and illustrated in Fig. 4.

The expectance $E[Z^2] = \langle Z^2 \rangle$ is computed directly from (114). The median $u = u_{\frac{1}{2}}$ is a special case of the α-fractile $u = u_\alpha$ defined as the solution of

$$F(u) = \alpha, \tag{C41}$$

where the cumulative distribution function $F(u)$ is found by integrating (99):

$$F(u) = \int_0^u p(t)dt = (n-1)(1-\gamma^2)^n \sum_{r=0}^{\infty} \left[\frac{(n+r-1)!}{(n-r)!r!}\right]^2 B_u(r+1,n-1) \gamma^{2r}. \tag{C42}$$

The equation (C41) is solved numerically by a Newton-Raphson process:

$$u_{k+1} = u_k - \frac{F(u_k) - \alpha}{p(u_k)} ; \tag{C43}$$

a sensible initial guess for the median will be

$$u_o = \gamma^2 \quad . \tag{C44}$$

The mode u = u* is the point (if it exists) where

$$p'(u) = 0 \quad . \tag{C45}$$

u* is the most probable outcome in a single realization. Using the derivation formula for hypergeometric functions

$$\frac{d}{dz} F(a,b;c;z) = \frac{ab}{c} F(a+1,b+1;c+1;z) \tag{C46}$$

we are led to solve

$$\varphi(u) \equiv n^2\gamma^2(1-u) \; F(n+1,n+1;2;\gamma^2 u)$$

$$- (n-2) \; F(n,n;1;\gamma^2 u) = 0 \quad . \tag{C47}$$

Like (C41), (C47) is solved by the Newton-Raphson method:

$$u_{k+1} = u_k - \frac{\varphi(u_k)}{\varphi'(u_k)} \tag{C48}$$

with the initial iterate this time chosen by the empirical rule

$$u_o = \min(\gamma^2 + \frac{1}{n}, \; 0.99) \quad . \tag{C49}$$

For the derivative of φ we find the following expression,

$$\varphi'(u) = \frac{1}{2}[n(n+1)]^2 \; \gamma^4(1-u) \; F(n+2,n+2;3;\gamma^2 u)$$

$$- n^2(n-1)\gamma^2 F(n+1,n+1;2;\gamma^2 u) \quad . \tag{C50}$$

For n < 2 the mode u* does not exist. For n > 2 the distribution is unimodal.

Sales distributors:
G.E.C. Gad Strøget
Vimmelskaftet 32
DK-1161 Copenhagen K, Denmark

Available on exchange from:
Risø Library, Risø National Laboratory,
P.O.Box 49, DK-4000 Roskilde, Denmark

ISBN 87-550-1193-4
ISSN 0106-2840